AD_____

Award Number: DAMD17-02-1-0427

TITLE: Energetics and structure prediction of the network of homo- and hetero-Oligomers Formed by the Transmembrane Domains of the ErbBReceptor Family of Proteins

PRINCIPAL INVESTIGATOR: Karen G. Fleming, Ph.D.

CONTRACTING ORGANIZATION: Johns Hopkins University
Baltimore, MD 21205

REPORT DATE: June 2006

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-06-2005	Final	1 Jun 2002 - 31 May 2006

4. TITLE AND SUBTITLE
Energetics and structure prediction of the network of homo- and hetero-Oligomers Formed by the Transmembrane Domains of the ErbBReceptor Family of Proteins

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-02-1-0427

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Karen G. Fleming, Ph.D.

E-Mail: karen.fleming@jhu.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Johns Hopkins University
Baltimore, MD 21205

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT:

The erbB/HER receptor proteins use a single transmembrane domain to mediate growth factor induced signal transduction across membranes. A complex network of both homo- and hetero-oligomeric interactions exists for the members of this receptor family. The goals of our research are to examine the role played by the transmembrane domain in this network of interactions. We have measured the energetics of homo- and hetero-oligomerization for the four human erbB transmembrane domains. We find all interactions to be modest in energy. We have developed a thermodynamic model to distinguish random interactions of proteins in micelles from preferential ones. We propose that the erbB transmembrane domains do not require strong preferential interactions in order to enhance the interaction propensity for a covalently bound soluble domain.

15. SUBJECT TERMS
Breast Cancer, Membrane Protein Association, Thermodynamics, Neu HER2

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	77	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

1. Introduction

The erbB/HER receptor family mediates growth factor induced signal transduction across membranes. A complex network of both homo- and hetero-oligomeric interactions exists for the four members of this receptor family [1]. Overexpression of the second receptor, erbB2, is observed in 20-30% of human breast cancers and correlates with poor clinical prognosis [2] [3]. The goals of our research are to examine the role played by the transmembrane domains of the erbB receptors in the network of receptor-receptor interactions.

2. Body

2.1. Overview

During the granting period we have completed all parts of all tasks described in the revised Statement of Work with the exception of Task 2, part c, which we have not been able to successfully complete by the end of the granting period. We are continuing to pursue this work and plan to submit additional grant applications to agencies that fund cancer research using the data we obtained in this combined IDEA and CAREER award as a basis for further investigation. The specific progress we have made, referenced to the approved, revised statement of work, is described below in the remainder of this section.

2.2. Generation of erbB2 transmembrane domain clones. (Task 1, Part A)

The goal of this task was to generate Staphylococcal nuclease (SN)[1] – transmembrane domain fusion proteins containing an erbB receptor transmembrane domain. This chimeric construct is similar to the SN-glycophorin A transmembrane domain construct, whose association energetics have been well studied [4].

Professor Mark Lemmon (University of Pennsylvania) generously provided us with a clone for the erbB2 transmembrane sequences in the form of a DNA plasmid. We used the polymerase chain reaction and standard cloning techniques to subclone this erbB2 transmembrane domain into our expression construct. In our construct, the transmembrane domain of erbB2 is expressed C-terminal to the SN reading frame. This is analogous to the chimeric construct that has been used previously for glycophorin A transmembrane domain [4]. The DNA sequences of all clones were verified by DNA sequencing using the facility at the Johns Hopkins Medical School.

We generated the SN-erbB2 expression clones for both the human and rat erbB2 transmembrane sequences. The rat sequence was used because a single mutation in the rat sequence (Valine to Glutamate at position 664, V664E) has been shown to promote receptor dimer formation, focus formation and cellular transformation in rat cell lines [5, 6]. It is unclear whether the corresponding mutation will have the same effect in the background of the human sequence [7]. To be able to test this, we thus cloned both

[1] A summary of all abbreviations used in this report is given on page 11.

TABLE 1: WILD TYPE ERBB2 TRANSMEMBRANE SEQUENCES THAT HAVE BEEN CLONED INTO THE STAPHYLOCOCCAL NUCLEASE EXPRESSION CASSETTE

Clone Name	AA Sequence of Transmembrane Region	Species	Note
SNerbB2	EPG**LTSIVSAVVGILLVVVLGVVFGILI**	Human	**WT**
SNneu	EPG**VTFIIATVVGVLLFLILVVVVGILI**	Rat	**WT (proto oncogene)**

The abbreviation we use for the clone name is given in the first column. The erbB2 sequences are shown in black in the second column. The grayscale "EPG" sequence at the N-terminus denotes where the SN linker region ends. The third column gives the species from which the sequence was derived, and the last column indicates that these sequences are the natural, wild-type sequences.

the rat and human sequences. The amino acid sequences of the wild type human and rat transmembrane domains of the clones we generated are shown in Table 1.

2.3. Generation of erbB2 transmembrane domain mutant clones. (Task 1, Part A)

The goal of this task was to generate Staphylococcal nuclease (SN) – transmembrane domain fusion proteins containing mutations in the erbB receptor transmembrane domain. This set of experiments will help to define the specificity of the protein-protein interaction that is encoded by the amino acid sequence. This chimeric construct is similar to the SN-glycophorin A transmembrane domain construct, whose association energetics have been well studied [4].

TABLE 2: MUTATED ERBB2 TRANSMEMBRANE SEQUENCES THAT HAVE BEEN CLONED INTO THE STAPHYLOCOCCAL NUCLEASE EXPRESSION CASSETTE.

Clone Name	Amino Acid Sequence of Transmembrane Region	Species	
SNerbB2neu	EPG**LTSIVSAVEGILLVVVLGVVFGILI**	Human	**Neu oncogene mutation mapped to human background**
SNerbB2N	EPG**LTSIVSAVVVILLVVVLGVVFGILI**	Human	**G to V in N-terminal motif**
SNerbB2C	EPG**LTSIVSAVVGILLVVVLGVVFVILI**	Human	**G to V in C-terminal motif**
SNerbB2NC	EPG**LTSIVSAVVVILLVVVLGVVFVILI**	Human	**G to V in both N-terminal and C-terminal motifs**
SNneuT	EPG**VTFIIATVEGVLLFLILVVVVGILI**	Rat	**Neu oncogene mutation V664E**

The abbreviation we use for the clone name is given in the first column. The erbB sequences are shown in black in the second column. The mutated residues are shown in red font that is one size larger. The grayscale "EPG" sequence at the N-terminus denotes where the SN linker region ends. The third column gives the species on which the sequence was based, and the last column indicates a narrative description of the mutants.

We have generated expression clones for several mutations in both the human and rat erbB2 transmembrane sequences. The amino acid sequences of these mutants are shown in Table 2. These mutant clones were generated using standard molecular biology protocols. All mutant sequences were verified by DNA sequencing at the facility located at the Johns Hopkins Medical School.

We chose to start with these mutations because they have been shown in the literature under various conditions to alter the oligomeric state as compared to the wild type sequence [8] [7]. Most recently, results from Lemmon and colleagues [7], demonstrated that these mutations caused changes in the ability of the erbB2 transmembrane domain to generate a result in a genetic assay that probes transmembrane helix-helix interactions. This, we hypothesized that these mutations would be most appropriate for the studies on the sequence specificity of the interaction.

2.4. Generation of erbB1, erbB3 and erbB4 transmembrane domain clones. (Task 1, Part A)

The goal of this task was to generate Staphylococcal nuclease (SN) – transmembrane domain fusion proteins containing each erbB receptor transmembrane domain. This chimeric construct is similar to the SN-glycophorin A transmembrane domain construct, whose association energetics have been well studied [4].

We ordered sense and antisense DNA oligonucleotides (~100 bp long each) whose sequences corresponding to the erbB1, erbB3 and erbB4 transmembrane domains. The sense and antisense strand of the DNA nucleotides for each erbB clone were combined in a 1:1 mole ratio, heated to 90 degrees and slow cooled to room temperature in order to form the double stranded DNA sequence encoding the transmembrane domain. This DNA insert was then mixed at several mole ratios with the pet11A-SN vector that had previously been restricted with SmaI and BamHI. Ligase was added, and the ligation reaction was allowed to proceed overnight. An aliquot of each ligation reaction was transformed into the DH5α cloning strain, and clones were screened for the presence of the correct insert using restriction digests. The DNA sequences of all clones were

TABLE 3: WILD TYPE ERBB TRANSMEMBRANE SEQUENCES THAT HAVE BEEN CLONED INTO THE STAPHYLOCOCCAL NUCLEASE EXPRESSION CASSETTE.

Clone Name	AA Sequence of Transmembrane Region	Species	Note
SNerbB1	EPG**SIATGMVGALLLLLVVALGIGLFMRRR**	Human	**WT**
SNerbB3	EPG**LTMALTVIAGLVVIFMMLGGTFLYWRGRR**	Human	**WT**
SNerbB4	EPG**LIAAGVIGGLFILVIVGLTFAVYVRRK**	Human	**WT**

The abbreviation we use for the clone name is given in the first column. The erbB sequences are shown in black in the second column. The grayscale “EPG” sequence at the N-terminus denotes where the SN linker region ends. The third column gives the species from which the sequence was derived, and the last column indicates that these sequences are the natural, wild-type (WT) sequences.

verified by DNA sequencing using the facility at the Johns Hopkins Medical School.

In the final construct, each erbB transmembrane domain is expressed C-terminal to the SN reading frame. Each of these constructs is analogous to the chimeric construct that has been used previously for glycophorin A transmembrane domain [4, 9-12]. The amino acid sequences of the transmembrane domains of the clones we generated in this second year are shown in Table 3.

2.5. Generation of erbB4 transmembrane domain mutant clones. (Task 1, Part A)

The goal of this task was to generate Staphylococcal nuclease (SN) – transmembrane domain fusion proteins containing mutations in the erbB receptor transmembrane domain. This set of experiments will help to define the specificity of the protein-protein interaction that is encoded by the amino acid sequence. This chimeric construct is similar to the SN-glycophorin A transmembrane domain construct, whose association energetics have been well studied [4, 9-12].

Using a QuickChange mutagenesis with the wild type SnerbB4 clones as a template, we have generated expression clones for several mutations in the human erbB4 transmembrane sequence. The amino acid sequences of these mutants are shown in Table 4.

We chose to start with these mutations because Lemmon and colleagues [7], demonstrated that these mutations caused changes in the ability of the erbB4 transmembrane domain to generate a result in a genetic assay that probes transmembrane helix-helix interactions in bacterial membranes. This, we hypothesized that these mutations would be most appropriate for the studies on the sequence

TABLE 4: MUTATED ERBB4 TRANSMEMBRANE SEQUENCES THAT HAVE BEEN CLONED INTO THE STAPHYLOCOCCAL NUCLEASE EXPRESSION CASSETTE.

Clone Name	Amino Acid Sequence of Transmembrane Region	Species	
SNerbB4N	EPG**LIAAGVIG<u>V</u>LFILVIVGLTFAVYVRRK**	Human	**G to V in N-terminal GxxxG motif**
SNerbB4C	EPG**LIAAGVIGGLFILVIVGLTF<u>V</u>VYVRRK**	Human	**A to V in C-terminal AxxxG motif**
SNerbB4NC	EPG**LIAAGVIG<u>V</u>LFILVIVGLTF<u>V</u>VYVRRK**	Human	**G to V in N-terminal GxxxG motif and A to V in C-terminal motif**

The abbreviation we use for the clone name is given in the first column. The erbB sequences are shown in black in the second column. The mutated residues are shown in red font that is one size larger. The grayscale “EPG” sequence at the N-terminus denotes where the SN linker region ends. The third column gives the species on which the sequence was based, and the last column indicates a narrative description of the mutants.

specificity of the interaction.

2.6. Expression and Purification of SN-Erb clones (Task 1, Parts B and C.)

The goal of this task was to express and purify the Staphylococcal nuclease (SN) – erbB TM domain fusion proteins.. This set of experiments will help to define the energetics and specificity of the protein-protein interaction that is encoded by the amino acid sequence.

This goal has been accomplished using our standard protocol for purifying chimeric proteins composed of Staphylococcal nuclease and transmembrane domains [13]. Briefly, the protein expression was induced in log-phase E. coli cells by the addition of 1 mM Isopropyl-D–thiogalactopyranoside (IPTG). After three hours of induction, the cells were harvested, lysed, and then the SN-erbB2 fusion protein was extracted using buffer containing 200 mM NaCl and 2% Thesit detergent. The induction of the SN-erbB fusion protein is determined by SDS-PAGE. Typical results are shown in Figure 1.

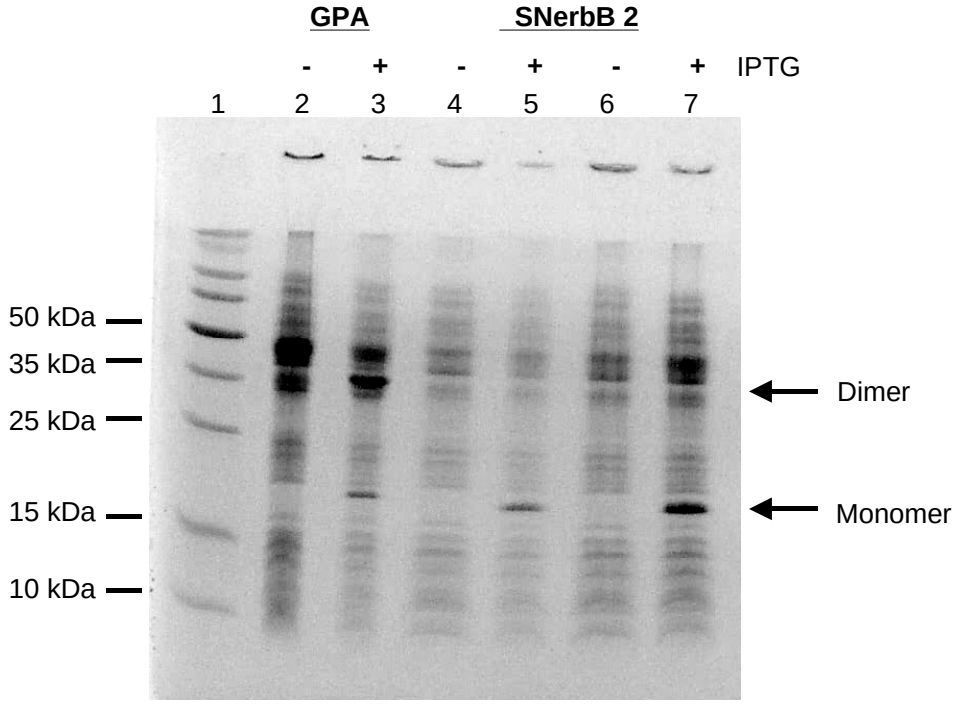


Figure 1: Expression Gel for SN-erbB 2. Whole cells were resuspended in SDS sample buffer and loaded onto a 20% polyacrylamide Phast gel. Lane 1 is the molecular weight markers, Lanes 2 and 3 are a control with the SN-glycophorin A transmembrane chimeric protein. Lanes 4 through 7 are two different SnerbB2 clones. The "-" lanes are uninduced cells, and the "+" lanes are induced cells that are expressing the SN fusion protein as indicated. The expected migration position of the monomeric and dimeric SN fusion proteins are indicated by arrows. The expression gel shows that the SN-GpA chimera expresses as a monomer and a dimer whereas the both SN-erbB2 proteins express as monomers.

2.7. Analyze purified SN-ErbB2 clones for association by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Task 2, Part A)

The goal of this task was to determine whether the erbB2 transmembrane domain drives strong protein-protein association of the SN-erbB fusion proteins. This SDS-PAGE technique was initially used to determine the sequence dependence of the glycophorin A transmembrane dimerization and is considered a method to detect strong dimerization.

We used 20% Phast Gels to determine the molecular weight of fusion proteins as measured by SDS-polyacrylamide gel electrophoresis (PAGE). The results for three of our SN-erbB2 variants are shown in Figure 2. The remaining variants show similar results. The SDS-PAGE assay shows that the proteins are monomeric as measured by this assay. This result suggests the conclusion that that their self-association propensity is weaker than that of the glycophorin A transmembrane domain [4, 11, 13]. Since the erbB protein is a receptor that is probably not constitutively dimeric, this result is not too surprising.

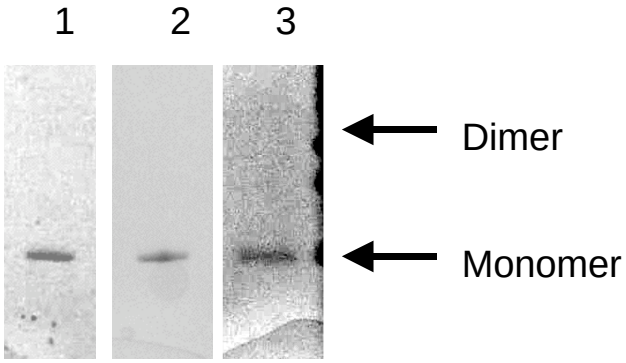


Figure 2: SDS-PAGE analysis to probe for strong dimerization driven by the erB2 transmembrane domain. Lane One shows the wild type SN-erbB2 protein. Lane Two shows the SnerbB2N protein. Lane Three shows the human SN-erbB2neu protein. All of these proteins migrated as monomers as measured by this assay. Expression tests like the type shown in Figure one were used to evaluate the SDS-PAGE migration of the remaining mutants. These were all found to be monomeric by this assay.

2.8. Expression and Purification of SNerb1, SNerb3 and SNerb4 wild type clones (Task 1, Parts B and C.)

The goal of this task was to express and purify the Staphylococcal nuclease (SN) – erbB transmembrane domain fusion proteins. The association energetics of the purified protein will then be analyzed using sedimentation equilibrium analytical ultracentrifugation. This set of experiments will help to define the energetics and specificity of the protein-protein interaction that is encoded by the amino acid sequence.

This goal has been accomplished for all proteins using our standard protocol for purifying chimeric proteins composed of Staphylococcal nuclease and transmembrane domains [13]. Briefly, the protein expression was induced in log-phase *E. coli* cells by the addition of 1 mM Isopropyl-D–thiogalactopyranoside (IPTG). After three hours of induction, the cells were harvested, lysed, and then the SN-erbB fusion protein was extracted using buffer containing 200 mM NaCl and 2% Thesit detergent. The induction of the SN-erbB fusion protein is determined by SDS-PAGE. Typical results are shown in Figure 3.

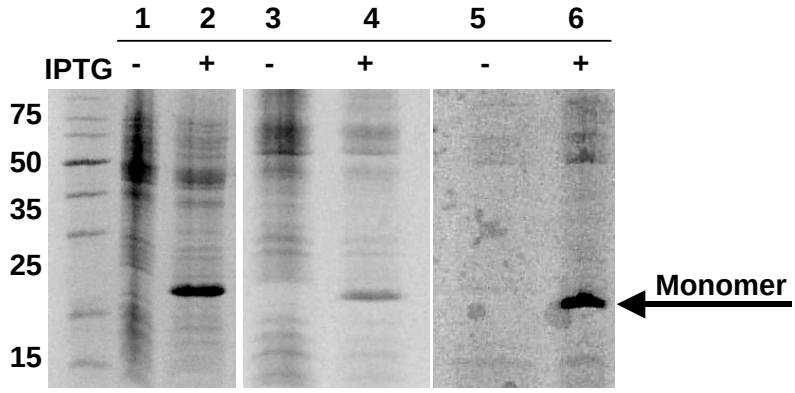


FIGURE 3: EXPRESSION GEL FOR SNERBB1, SNERBB3 AND SNERBB4. Whole cells were resuspended in SDS sample buffer and loaded onto a 20% polyacrylamide Phast gel. Lane 0 is the molecular weight markers, Lanes 1 and 2 are SnerbB1, lanes 3 and 4 are SNerbB3, lanes 5 and 6 are SNerbB4. The "-" lanes are uninduced cells, and the "+" lanes are induced cells that are expressing the SN fusion protein as indicated. The expected position of the monomeric SN fusion proteins are indicated by the arrow. The expression gel shows that all of the SN-erbB proteins express as monomers.

2.9. Analyze purified SN-ErbB clones for association by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Task 2, Part A)

The goal of this task was to determine whether the erbB transmembrane domains drive strong protein-protein association of the SN-erbB fusion proteins. This SDS-PAGE technique was initially used to determine the sequence dependence of the glycophorin A transmembrane dimerization and is considered as a method to detect strong dimerization.

We used 20% Phast Gels to determine the molecular weight of fusion proteins as measured by SDS-polyacrylamide gel electrophoresis (PAGE). The results for all of the wild type SNerbB variants are shown in Figure 4. The SDS-PAGE assay shows that the purified proteins are monomeric as measured by this assay. This result suggests the conclusion that that their self-association propensity is weaker than that of the glycophorin A transmembrane domain [4, 11, 13]. Since all of the erbB proteins are transmembrane signaling receptors that are probably not constitutively dimeric, this result is not too surprising.

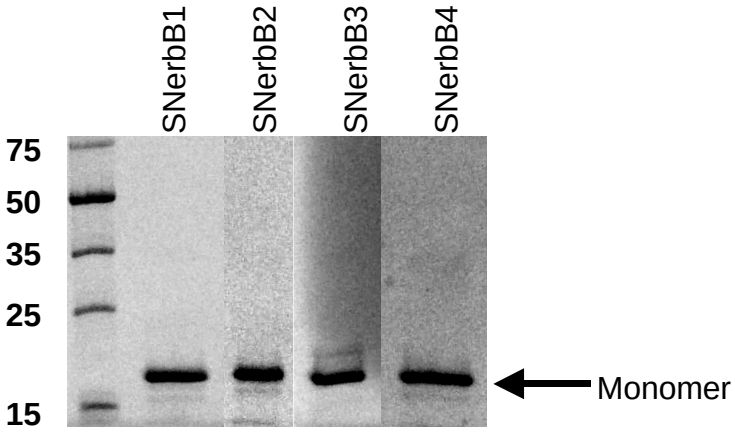


FIGURE 4: SDS-PAGE ANALYSIS OF PURIFIED SNERBB PROTEINS TO PROBE FOR STRONG DIMERIZATION DRIVEN BY THE ERBB TRANSMEMBRANE DOMAIN.

Lane One shows the molecular weight markers. Lanes 2-5 show the wild type SN-erbB1, SN-erbB2, SN-erbB3 and SN-erbB4 proteins. All of these proteins migrated as monomers as measured by this assay.

2.10. Measurement of the free energy of association for the SNerbB1, SNerbB2, SNerbB3 and SNerbB4 transmembrane domain using sedimentation equilibrium analytical ultracentrifugation. (Task 2, Part B)

To further test for the self-association propensity of the erbB transmembrane domains, we carried out sedimentation equilibrium analytical ultracentrifugation experiments. Sedimentation equilibrium analytical ultracentrifugation allows detection of protein-protein interactions over a range of association strengths. In particular, we have detected significant dimer populations of the SN-GpA mutant proteins even when these proteins are monomeric as measured by SDS-PAGE [4, 13]. In addition, unlike SDS-PAGE, sedimentation equilibrium experiments can provide a direct measure of the equilibrium constant for a self-association reaction.

We followed a standard sedimentation equilibrium protocol that we developed and that has been described in detail in the literature [4, 13, 14]. Data were collected at three different speeds (20000, 24500 and 30000) starting with samples at three different initial protein concentrations. This sedimentation equilibrium experiments were repeated in different detergent micelle concentrations.

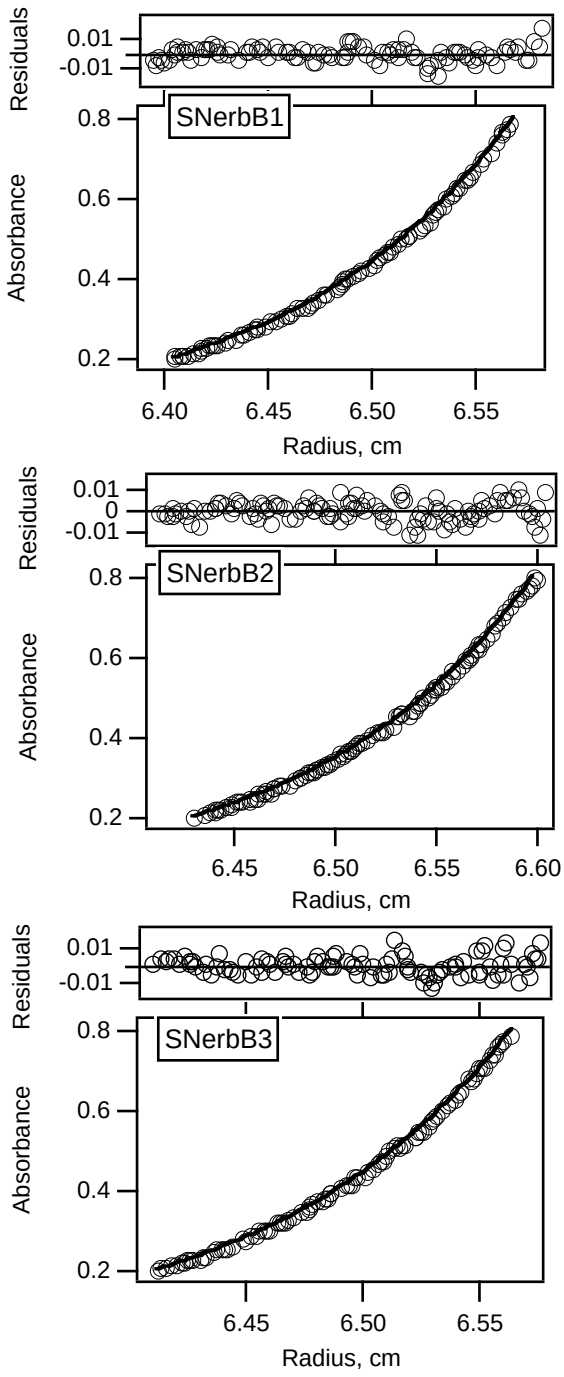
Our major finding is that all SNerbB transmembrane domains self-associate quite weakly in detergent micelles. The data to support this conclusion is shown on the next page.

2.11. SNerbB1, SNerbB2 and SNerbB3 are monomeric in 23 mM C8E5.

In 23 mM C_8E_5 we find that the SNerbB1, SNerbB2 and SNerbB3 transmembrane domains are well described as a single species in solution with a molecular weight corresponding to that of the monomeric protein. A representative fit for each protein in 23 mM C_8E_5 is shown in Figure 5. At this same concentration of C_8E_5, the WT glycophorin A transmembrane domain is > 90% dimeric.

FIGURE 5: THE SNERBB1, SNERBB2, AND SNERBB3 PROTEINS ARE MONOMERIC IN 23 MM C_8E_5.

Representative data sets from sedimentation equilibrium analytical ultracentrifugation experiments are shown. The protein distribution was monitored by absorbance at 280nm. Raw data is shown as open circles; the solid lines represent global fits for single ideal species. This fit returns the best estimate molecular weight value corresponding to the monomeric protein. The upper panel for each data set shows the residuals for the fit. The residuals for the single ideal fit are small and random for these three proteins.



2.12. SNerbB4 is not monomeric in 23 mM C_8E_5.

In contrast to the previous erbB transmembrane domains, the SNerbB4 transmembrane domain is not monomeric in 23 mM C_8E_5. Shown in Figure 6, an attempt to fit this protein to a single species yields a poor fit as well as a molecular weight estimate that is greater than that of the monomeric protein.

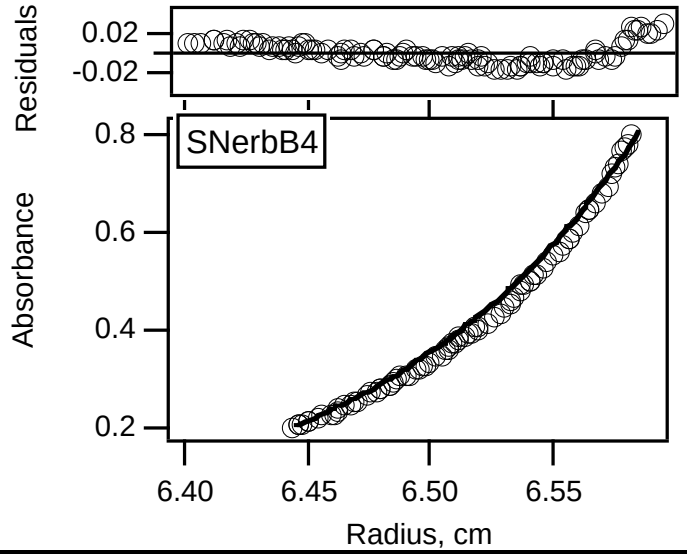


FIGURE 6. SNERBB4 IS NOT MONOMERIC IN 23 MM C_8E_5 MICELLES. Shown is a representative dataset reflecting the a global fit of three rotor speeds and three initial protein concentrations. The SNerbB4 protein distribution was monitored by absorbance at 280nm. Raw data is shown as open circles; the solid line represents the global fit for single ideal species. The upper panel for the data set shows the residuals for the fit, which are calculated as the difference between the data and the fit at each point. These residuals for the single ideal fit applied to SNerbB4 are not randomly distributed about zero, but rather have a "smile" shape". In addition, the magnitude of the residuals is twice as large as the magnitude for the residuals of the fits shown in Figure 5.

To determine the molecular weight value of the higher order species present, we attempted to fit this SNerbB4 23 mM data to models containing dimers, trimers, tetramers, as well as various combinations of species. None of these resulted in a unique fit to the data. Even though the average molecular weight value was higher than that of monomer, we think that the concentration of higher order species was still low in 23 mM C_8E_5, and this fact makes it difficult to firmly establish a molecular weight.

2.13. SNerbB4 associates more strongly in 11 mM C_8E_5.

To determine the nature of the higher molecular weight species formed by self-association of the SNerbB4 transmembrane domain, we carried out experiments in a lower detergent concentration. We chose 11 mM C_8E_5 because it is approximately three

times the critical micelle concentration for this detergent, and it is about half the concentration we used in our initial experiments. Decreasing the overall concentration of detergent has the effect of concentrating the protein in the micellar phase. As expected SNerbB4 self-associates to a greater extent in 11 mM compared to 23 mM C_8E_5.

In our first experiment, we found that a model containing only Monomer and Dimer species was a poor description of the data. This is shown in Figure 7.

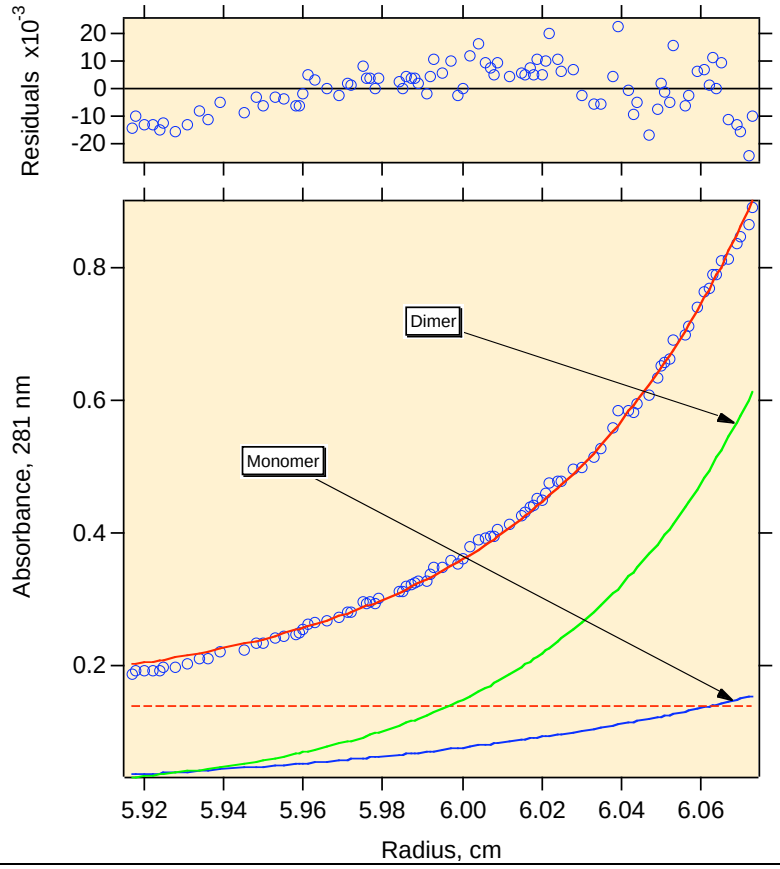


FIGURE 7. A SIMPLE MONOMER-DIMER FIT IS A POOR DESCRIPTION SNERBB4 SEDIMENTATION EQUILIBRIUM DATA. The lower panel shows a representative fit to a Monomer-Dimer model. The Monomer and Dimer curves along with the baseline (broken horizontal line) are summed to give rise to the solid line fit through the data. The upper panel shows the residuals of the fit. These show systematic deviations from zero and are non-random. All 9 of the data sets in the global monomer-dimer fit are poorly described.

Upon considering alternative mathematical descriptions of the data, we found that the simplest model to globally describe the sedimentation equilibrium data was the Monomer-Trimer model (see Table 5). Because this is an unexpected and completely unprecedented result from a functional point of view, we repeated the experiment four separate times using different protein preparations and different detergent stocks. (Each of these experiments takes about 3 weeks to complete). The summary of the global fits

are given in Table 5 in the form of the values for the square root of the variance (SRV). Normally, the model with the lowest SRV value is taken as the best fit description of the data.

TABLE 5: INITIAL SEDIMENTATION EQUILIBRIUM ANALYSIS MIGHT INDICATE THAT SNERBB4 HOMOTRIMERIZES IN C_8E_5 MICELLES. Shown are the values for the square root of variance (SRV) for each fit for four independent experiments on the SNerbB4 protein. Normally, the simplest model with the lowest value for the SRV is the best description of the data.

Model applied during global fit of SNerbB4 data in 11 mM C_8E_5	Experiment Number				
	1	2	3	4	
Single Ideal*	8.66	9.32	8.52	8.46	
Monomer-Dimer	7.72	8.31	7.34	6.81	
Monomer-Trimer	6.66	7.23	6.98	5.64	←Lowest value in all cases
Monomer-Tetramer	8.34	8.94	8.73	7.64	

*In all cases, the single, ideal fit returned a molecular weight value higher than the calculated molecular weight of the monomer. This result indicates self-association in the analytical ultracentrifuge. For clarity, the SRV values given here are the original ones that have been multiplied by a factor of 1000.

In every case, the global fits to the data with two species showed the lowest SRV value for the Monomer-Trimer fit when compared to Monomer-Dimer and Monomer-Tetramer fits. We noted that the statistics of the Monomer-Trimer fit were within the criteria for acceptance, however, in the sedimentation equilibrium field, it is well known that numerically a Monomer-Trimer fit can sometimes be a good mathematical description of sedimentation equlibrium data in cases when there is mixture of three species that include dimer along with a small amount of higher order oligomer, which is typically tetramer. Since the formation of dimer is more consistent with the functional data for the erbB proteins, we therefore tested additional Monomer-Dimer models containing three species. For the same data, these are shown in Table 6 below.

TABLE 6: SEDIMENTATION EQUILIBRIUM ANALYSIS OF SNERBB4 PROTEINS THAT INCLUDE THREE SPECIES. Shown are the values for the square root of variance (SRV) for each fit for four independent experiments on the SNerbB4 protein.

Model applied during global fit of SNerbB4 data in 11 mM C8E5	Experiment Number				
	1	2	3	4	
Monomer-Dimer-Trimer	6.60	7.16	6.76	5.56	
Monomer-Dimer-Tetramer	6.70	7.19	6.82	5.68	
Monomer-Dimer-Sep Trimer	5.36	6.08	6.36	4.89	←Lowest value in all cases

Indeed, a fit with three species including dimer results in lower SRV values. In all experiments, the Monomer-Dimer Sep Trimer model is the clearly the best description of the data. The presence of these three species was confirmed by crosslinking experiments shown on the next page.

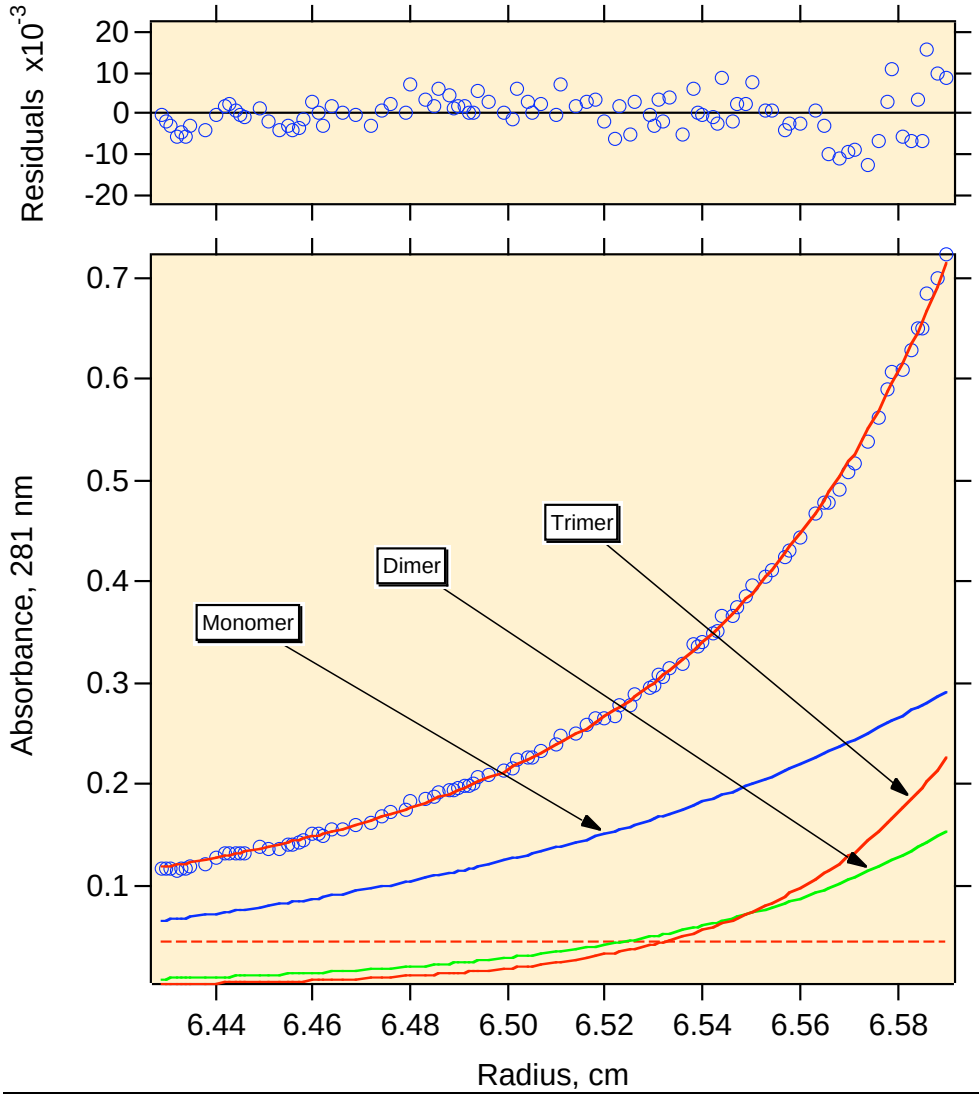


FIGURE 8. A MONOMER-DIMER-TRIMER FIT IS AN ACCEPTABLE DESCRIPTION SNERBB4 SEDIMENTATION EQUILIBRIUM DATA. The lower panel shows a representative fit to a Monomer-Dimer-Trimer model. The Monomer, Dimer and Trimer curves along with the baseline (broken horizontal line) are summed to give rise to the solid line fit through the data. The upper panel shows the residuals of the fit. In contrast to figure 7 these show more random deviations about zero, which is expected for a good fit.

2.14. Cross-linking of SNerbB4 confirms that it preferentially forms dimers and trimers.

To independently determine the preferred association species for the SNerbB4 protein, we carried out crosslinking experiments followed by SDS-PAGE analysis. The crosslinking experiments were carried out using the same protein and detergent concentrations as the sedimentation equilibrium species. The crosslinking results for SNerbB4 are shown in Figure 9. These experiments independently demonstrated that SNerbB4 does indeed form dimers as well as small amounts of trimers.

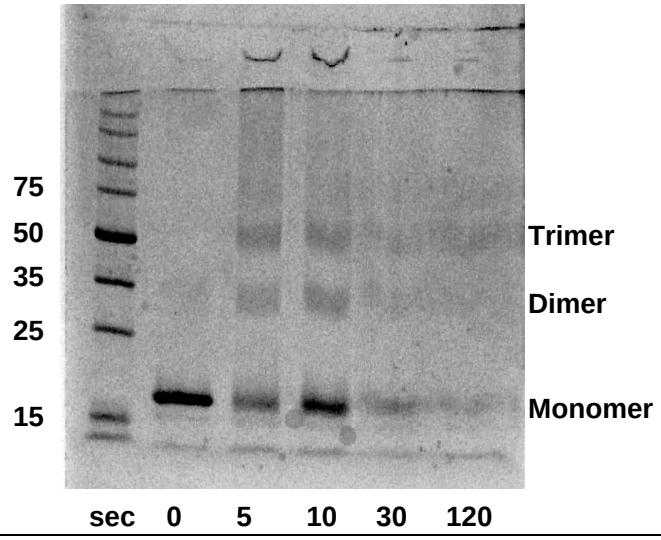


FIGURE 9: CROSSLINKING CONFIRMS THAT SNERBB4 FORMS DIMERS AND TRIMERS.

Shown is SDS-PAGE analysis of glutaraldehyde cross-linked SNerbB4 as a function of time with the crosslinker. The proteins were cross-linked in 11mM C_8E_5 at 43μM with 0.1%glutaraldehyde and run on a 12.5% acrylamide Phast Gel. Lane 1, Molecular weight markers; Lanes 2-5 show the time course of crosslinking. Bands can be seen for both dimers and trimers; traces of higher order aggregates can also be seen. Note that after ~30 seconds, all of the SNerbB4 protein crosslinks as large aggregates that do not enter the gel.

The crosslinking results justify the use of dimeric and trimeric species in the analytical ultracentrifugation analysis. From the combination of the crosslinking and sedimentation equilibrium data, we conclude that a model containing Monomeric, Dimeric and Trimeric species would be required to describe the data. The presence of a large molecular weight aggregate would not be observed in the ultracentrifuge at the speeds we are using. If this large molecular weight species formed irreversibly in the absence of the crosslinking agent, the sedimentation equilibrium experiment would show a loss of detectable protein over the course of the centrifuge experiment. We do not observe a loss of mass in the ultracentrifuge, and since the time to sedimentation equilibrium (for all three speeds) is approximately 72 hours, we conclude that this large aggregate is not appreciably populated in the absence of the crosslinking agent.

The statistics of the global fits to the sedimentation equilibrium data suggest that the best model for association is the Monomer-Dimer-Separate Trimer. Thermodynamically, this model suggests that the monomer and dimer are reversibly associating whereas the trimer species is variable between experiments, and it is not described by a thermodynamically meaningful dimer-trimer association constant. The apparent monomer dimer dissociation constants and free energies of association for the four independent experiments are given Table 7 below:

TABLE 7: THERMODYNAMIC PARAMETERS FOR SNERBB4 DIMERIZATION. Shown are the best fit estimates for the association and dissociation equilibrium constants for the Monomer-Dimer-Separate Trimer global fits for four independent experiments carried out on the SNerbB4 protein. The apparent free energy of dimerization in 11 mM C_8E_5 calculated from the association constant determined at 25 degrees C.

Exp Number	K_a (Molar^{-1})	K_d (Molar)	ΔG_{app} (kcal mol^{-1})
1	23075	4.33E-05	**-5.948**
2	22823	4.38E-05	**-5.942**
3	22816	4.38E-05	**-5.942**
4	14703	6.8E-05	**-5.681**
Average	20854	4.8E-05	**-5.89**
Standard Deviation			**±0.13**

2.15. At lower detergent concentrations, the SNerbB1, SNerbB2, and SNerbB3 transmembrane domains self-associate.

Since lowering the detergent concentration from 23 mM to 11 mM had the effect of populating the dimeric form of SNerbB4, we carried out experiments on the other erbB transmembrane domains to determine whether they too had the potential to form higher order species. Indeed, in our preliminary experiments described last year, we did find that SNerbB2 could form dimers in 11 mM C_8E_5.

Our sedimentation equilibrium experiments showed that a single species monomeric fit is no longer a good description of the behavior of the SNerbB1, SnerbB2 and SNerbB3 transmembrane domains when the detergent concentration is lowered from 23 mM to 11 mM C_8E_5. A good model-independent parameter for expressing the extent of association is the value for sigma derived from the single species fit. Sigma is defined by Yphantis [15] as:

$$\sigma = \frac{M(1-\bar{v}\rho)\omega^2}{RT} \qquad (1)$$

where M is the molecular weight, $\bar{v}$ is the partial specific volume of the protein (ml g^{-1}), ρ is the solvent density (g ml^{-1}), ω is the rotor speed (rad sec^{-1}), R is the universal gas constant, and T is the temperature (°K). It is standard to calculate the partial specific volume from the amino acid composition, so it is a known value. For a single species fit, the value of sigma will represent an approximate measure of the weight average molecular weight of the proteins in solution. If the proteins are monomeric, then the single species fit will return an experimental value for sigma, σ_{Exp}, that is equal to the value of sigma one would calculate from knowing the amino acid sequence for the protein, σ_{Calc}. Thus a ratio of $\left(\frac{\sigma_{Exp}}{\sigma_{Calc}}\right)$ that is greater than one is an indication of the presence of higher order species. The precision of the sedimentation equilibrium experiment is normally ±5-10%. Thus, a σ_{Exp} that exceeds the σ_{Calc} by more than ten percent suggests self-association. Table 8 shows the sigma ratios for all four SNerbB proteins in both the 23 mM C_8E_5 and 11 mM C_8E_5 experiments.

TABLES 8. SIGMA VALUES FOR SINGLE SPECIES FITS OF THE FOUR SNERBB PROTEINS IN TWO DETERGENT MICELLE CONCENTRATIONS.

		23 mM C_8E_5		11 mM C_8E_5	
SNerbB	Calculated Sigma for Monomer σ_{Calc}	Experimentally Determined Sigma* σ_{Exp}	Sigma Ratio $\left(\frac{\sigma_{Exp}}{\sigma_{Calc}}\right)$	Experimentally Determined Sigma σ_{Exp}	Sigma Ratio $\left(\frac{\sigma_{Exp}}{\sigma_{Calc}}\right)$
1	0.9575	1.0036	1.0481	1.3657	1.4263
2	0.9287	0.9844	1.0600	1.2943	1.3937
3	0.9841	1.0581	1.0751	1.3593	1.3813
4	0.9533	1.1088	1.16311	1.6260	1.7057

*The experimental sigma values in the table represent the average of at least two independent measurements.

Consistent with Figure 5 in section 2.10, SNerbB1, SNerbB2 and SNerbB3 have sigma ratios of ~1.0 (±0.1) in 23 mM C_8E_5 demonstrating that these three proteins are monomeric under these conditions. In contrast, in detergent concentrations of 11 mM, the SNerbB proteins show sigma values greater than monomer, and SNerbB4 shows the greatest amount of self-association in both 11 and 23 mM C_8E_5.

2.16. Trends in the extent of self-association in micelles are similar to those observed in membranes.

Since the sigma value is an indication of the weight average molecular weight, which is a reflection of the extent of self-association, we compared the 11 mM sigma values to the extent of association that has been determined by Lemmon and colleagues [7] using the TOXCAT genetic assay in bacterial membranes. The TOXCAT assay uses a construct encoding the periplasmic protein, MBP, fused to a query

transmembrane domain sequence, fused to the transcription domain of the ToxR transcription domain from Vibrio cholera. A cartoon for this assay is shown in Figure 10. If the transmembrane sequence drives self-association of the fusion protein, then the ToxR transcription domains will be brought into proximity with each other at which point the ToxR-ToxR dimer (or higher-order oligomer) can bind to the cholera toxin promoter and activate transcription of a reporter gene. Using this assay, Lemmon and colleagues tested the transmembrane domains of all four human erbB receptors for their ability to produce CAT activity. Since the assay is a relative one, each measurement is compared to a positive control, the wild-type glycophorin A transmembrane domain.

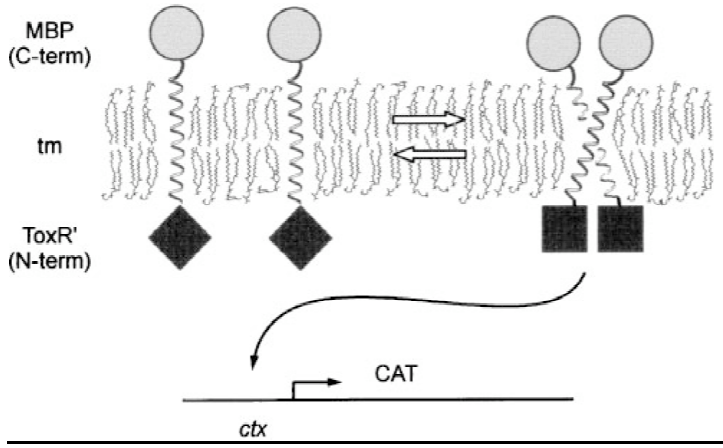
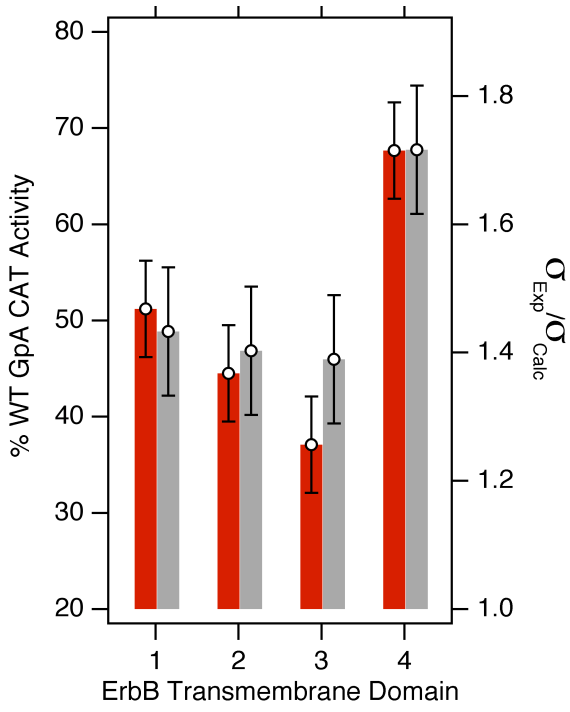


FIGURE 10: SCHEMATIC DEPICTING HOW THE GENETIC SCREENS ARE THOUGHT TO MEASURE HELIX-HELIX ASSOCIATION IN BACTERIAL MEMBRANES. If the TM segments drive self-association then the ToxR' domains will dimerize and become active for binding to the *ctx* promoter region.

Figure 11 shows a bar graph with the relative TOXCAT signals for each of the erbB transmembrane domains expressed as a percentage of the WT glycophorin A signal. Next to each bar is shown the sigma ratios that we determined using sedimentation equilibrium. In both types of experiment, the erbB4 transmembrane domain shows the greatest extent of association, and the other three erbB

FIGURE 11. TRENDS IN AVERAGE MOLECULAR WEIGHT ARE CONSISTENT WITH TRENDS IN THE ASSOCIATION PROPENSITY MEASURED IN BIOLOGICAL MEMBRANES USING THE TOXCAT GENETIC ASSAY. Shown in the dark red bars are the CAT activity expressed as a percentage of the glycophorin A (GpA) transmembrane domain positive control. Shown in the lighter grey bars are the mean sigma values for each SNerbB protein determined by sedimentation equilibrium. In both types of experiment, the erbB4 transmembrane domain shows the greatest extent of association, and the other three erbB transmembrane domains show extents of association similar to each other but less than that of the erbB4 transmembrane domain.



transmembrane domains show extents of association similar to each other but less than that of the erbB4 transmembrane domain.

2.17. Crosslinking experiments with SNerbB1, SNerbB2 and SNerbB3

We carried out crosslinking experiments to determine the nature of the nature of the higher molecular weight species that are formed by the SNerbB1, SNerbB2 and SNErbB3 proteins. The cross-linked species were resolved by SDS-PAGE, and the time courses for these three proteins are shown in Figures 12-14.

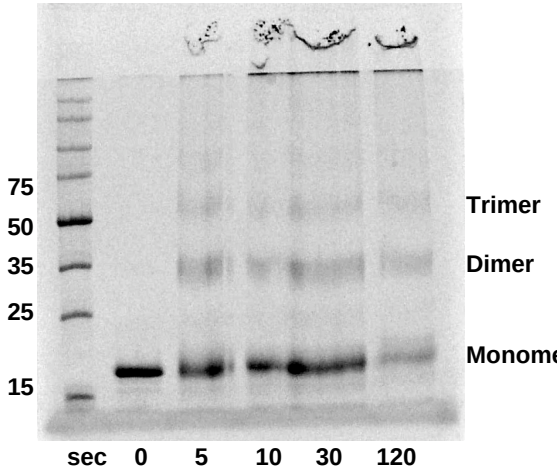


FIGURE 12: CROSSLINKING OF SNERBB1. Shown is SDS-PAGE analysis of glutaraldehyde cross-linked SNerbB1 as a function of time. SNerbB1 was cross-linked in 11mM C8E5 at 43μM with 0.1%glutaraldehyde and run on a 12.5% acrylamide Phast Gel. Lane 1, Molecular weight markers; Lanes 2-5 show the time course of crosslinking. Bands can be seen for both dimers and trimers; traces of higher order aggregates can also be seen.

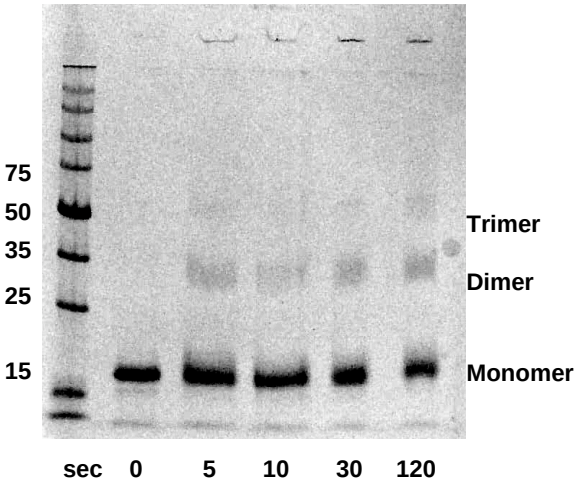


FIGURE 13: CROSSLINKING OF SNERBB2. Shown is SDS-PAGE analysis of glutaraldehyde cross-linked SNerbB2 as a function of time with the crosslinker. The proteins were cross-linked in 11mM C8E5 at 43μM with 0.1%glutaraldehyde and run on a 12.5% acrylamide Phast Gel. Lane markers are the same as Fig. 10.

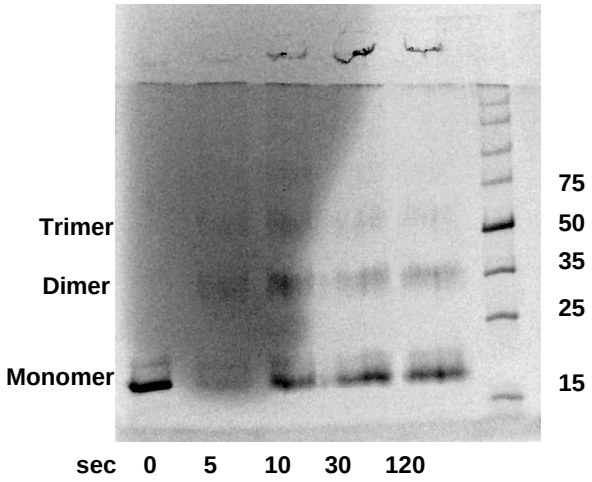


FIGURE 14: CROSSLINKING OF SNERBB3. Shown is SDS-PAGE analysis of glutaraldehyde cross-linked SNerbB2 as a function of time with the crosslinker. SNerbB3 was cross-linked in 11mM C_8E_5 at 43μM with 0.1%glutaraldehyde and run on a 12.5% acrylamide Phast Gel. Lane 1, Molecular weight markers; Lanes 2-5 show the time course of crosslinking. Bands can be seen for both dimers and trimers; traces of higher order aggregates can also be seen.

2.18. Species analysis of sedimentation equilibrium data

TABLE 9: SEDIMENTATION EQUILIBRIUM ANALYSIS OF SNERBB1 PROTEINS THAT INCLUDE THREE SPECIES. Shown are the values for the square root of variance (SRV) for each fit for four independent experiments on the SNerbB1 protein. (A combination of the lowest SRV value and the simplest model indicates the best fit).

	Experiment Number	
Model Tested	1	2
Single Ideal	7.93	7.96
Monomer-Dimer	7.41	6.55
Monomer-Trimer	7.78	6.01
Monomer-Tetramer	8.96	6.68
Monomer-Dimer-Trimer	7.39	5.99
Monomer-Dimer-Tetramer	7.41	6.01
Monomer-Dimer-Sep Trimer	7.03/6.50	5.4

TABLE 10: SEDIMENTATION EQUILIBRIUM ANALYSIS OF SNERBB2 PROTEINS THAT INCLUDE THREE SPECIES. Shown are the values for the square root of variance (SRV) for each fit for four independent experiments on the SNerbB1 protein. (A combination of the lowest SRV value and the simplest model indicates the best fit).

	Experiment Number						
Model Tested	1	2	3	4	5	6	7(KKR)
Single Ideal	7.47	8.1	8.14		6.48		8.9
Monomer-Dimer	6.25	6.96	6.71		5.75		7.44
Monomer-Trimer	6.02	6.43	5.89		5.45		6.71

Monomer-Tetramer	7.15	7.02	6.59		5.6		7.94
Monomer-Dimer-Trimer	4.91	4.33	5.91	6.43	5.89	5.45	6.66
Monomer-Dimer-Tetramer	5.98	6.47	5.93		5.45		6.76
Monomer-Dimer-Sep Trimer	4.42	4.25	5.15	5.76	5.21	4.87	5.94

TABLE 11: SEDIMENTATION EQUILIBRIUM ANALYSIS OF SNERBB3 PROTEINS THAT INCLUDE THREE SPECIES. Shown are the values for the square root of variance (SRV) for each fit for four independent experiments on the SNerbB1 protein. (A combination of the lowest SRV value and the simplest model indicates the best fit).

	Experiment Number		
Model Tested	1	2	3
Single Ideal	6.78	9.18	5.56
Monomer-Dimer	6.28	7.95	4.55
Monomer-Trimer	6.37	7.24	4.27
Monomer-Tetramer	6.77	7.95	4.72
Monomer-Dimer-Trimer	6.22	7.22	4.25
Monomer-Dimer-Tetramer	6.21	7.2	4.26
Monomer-Dimer-Sep Trimer	5.5	6.6	6.57/4.14

2.19. Assay for heterodimerization of erbB transmembrane domains. (Task 2, Parts A & B)

2.19.1. SDS-PAGE Analysis of Potential Heterodimers.

The goal of this task was to determine whether the erbB transmembrane domains would mediate heterodimer formation. Our initial screen was carried out using

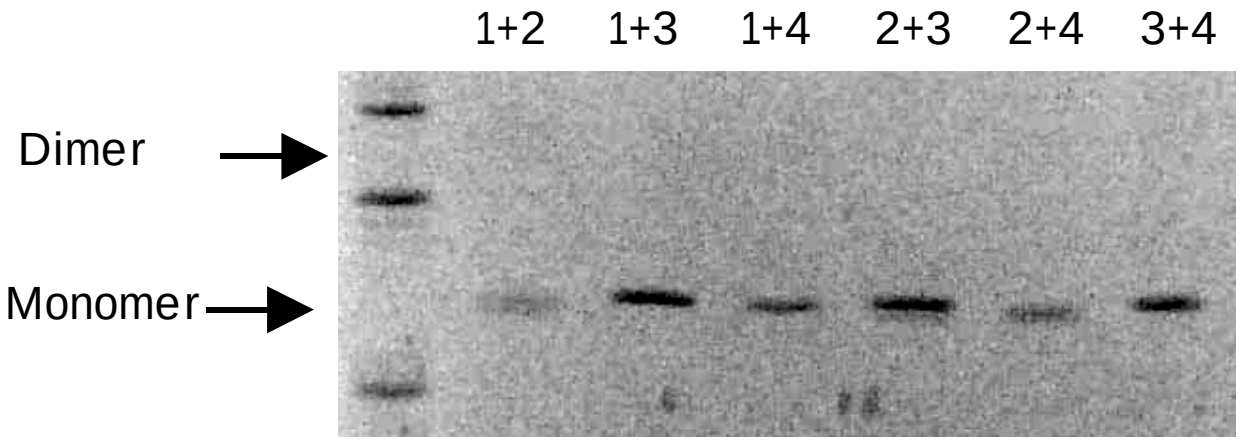


Figure 15: SDS-PAGE analysis of SNerbB mixing experiments using a 20% acrylamide pHast gel. All samples migrated to a position consistent with monomers in solution. Lane 1, Markers; Lane 2, SNerbB1/B2; Lane 3, SNerbB1/B3; Lane 4, SNerbB1/B4; Lane 5, SNerbB2/B3; Lane 6, SNerbB2/B4; Lane 7, SNerbB3/B4.

SDS-PAGE analysis, and we considered all six possible heterodimer configurations. SDS-PAGE analysis can detect the presence of strong transmembrane helix-helix interactions, in which case the proteins would migrate in oligomeric forms. The results are shown in Figure 15. As can be seen in Figure 15, we found that no pair of erbB transmembrane domains dimerized strongly.

2.19.2. Sedimentation Equilibrium Analysis of Potential Heterodimers

We also used sedimentation equilibrium to attempt to detect the presence of modest interactions mediated between SNerbB2 and SNerbB3. We chose this pair because the erbB2/erbB3heterodimer is the most potent signaling complex [16, 17], and only weak heterointeractions can be detected between the soluble ligand-binding domains [18], so we hypothesized that if transmembrane helix–helix interactions were important for stabilizing any of the hetero-complexes it would be the erbB2/ erbB3 complex. Our results are shown in Figure 16. They demonstrate that an equimolar amount of SNerbB2 and SNerbB3 at 23 mMC_8E_5 yielded a molecular mass consistent with protein monomers in solution.

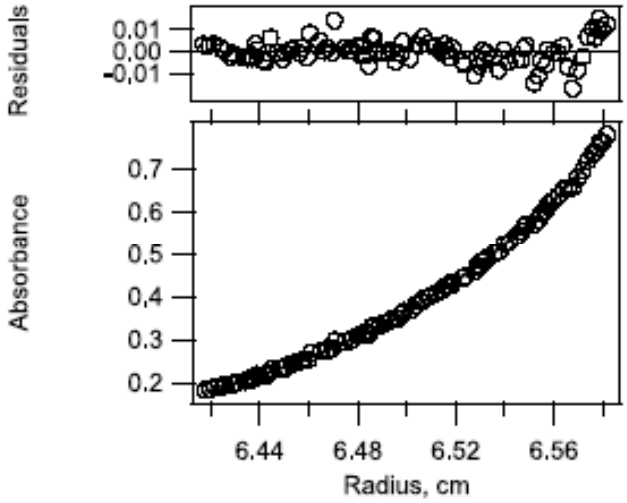


Figure 16: A representative sedimentation equilibrium data set for a mixture of equimolar amounts of SNerbB2 and SNerbB3 fit to a single ideal species. The raw data are represented as open circles and the continuous line represents the fit and the upper panel shows the residuals for the fit. The data are well described by a single ideal species with a molecular mass of 21,273 kDa, which is equal to the average of the two monomers, SNerbB2 (20,876 kDa) and SNerbB3 (21,669 kDa).

2.20. A Theoretical Formalism to Explain the Modest Free Energies of Self-Association that we observed. (Task 3)

2.20.1. Theoretical formalism.

The goal of this research was to determine whether the modest free energies of self-association we observed for the erbB transmembrane domains represented a thermodynamically meaningful preferential association of the population.

The crux of the issue lies in the fact that the sedimentation experiments were carried out at protein:detergent ratios much higher than were previously used for glycophorin A [12] in order to observe SN-erbB interactions. At the highest SN-erbB protein concentration the number of protein molecules approached the number of micelles. Since the fundamental measurement in the sedimentation equilibrium under the current experimental conditions is the protein mass, two protein molecules that occupy the same micelle will sediment together and be measured as a dimer whether or not there is a preferential interaction between them. This co-partitioning phenomenon has been referred to in the literature as "artifactual togetherness" [19] and should only occur at high protein concentrations coupled with low detergent micelle concentrations. We hypothesize that this co-partitioning is due to a type of excluded volume effect and that these conditions can be thought of as thermodynamically nonideal. A similar excluded volume effect has been indicated to occur in soluble proteins at high protein concentrations or in the presence of small molecules that induce crowding conditions [20]. To determine the extent of thermodynamically meaningful preferential interactions between the SN-CCK4 transmembrane proteins it is first necessary to distinguish them from the statistical protein interactions induced by the limiting solvent conditions.

To develop a model describing the distribution of non-interacting proteins in micelles, we employed a statistical approach that considers distribution of proteins in micelles in terms of the random occupancy of placing balls in bins, where the protein molecules represent balls and the individual micelles represent bins. Using this formalism, the random occupancy of any particular micelle by any particular protein molecule can be described as a Bernoulli event (*e.g.* either the micelle contains that protein or not), which can be described by a binomial distribution:

$$\Pr[X_i = k] = \binom{m}{k} p^k (1-p)^{m-k} \qquad (2)$$

where X_i is the random variable that counts the number of proteins in a particular micelle, i, k is the number of proteins occupying a micelle (*e.g.* the "oligomeric state"), m is the number of protein molecules, p is the probability of each Bernouilli trial, which equals $1/n$ and n is the number of micelles. When the number of proteins and the number of micelles both exceed the allowed occupancy of a particular micelle, k, this binomial distribution is approximated by the Poisson distribution:

$$\lim(m,n >> k) : \Pr[X_i = k] \approx \frac{1}{k!}\left(\frac{m}{n}\right)^k e^{-m/n} \qquad (3)$$

To consider the statistical protein distribution, we write the partition function, P, as the sum all those probabilities for which $k > 0$. (*e.g.* For this purpose, we are not

interested in the distribution of “empty” micelles). To convert each probability to a relative concentration, each term is multiplied by the number of proteins that the probability represents.

$$P = \sum_{k>0} k \Pr[X_i = k] \quad (4)$$

The fraction of any particular species can then be calculated as a function of the mole fraction protein (m/n) in the micellar phase from this partition function by dividing its contribution by the sum of all species. For example the fraction dimer can be calculated as:

$$f_{Dimer} = \frac{2 \Pr[X_i = 2]}{\sum_{k>0} k \Pr[X_i = k]} \quad (5)$$

2.20.2. Application of this formalism to the SN-erbB Distributions

The exact form of the partition function and the distribution of species will depend on the allowed values of k in the model. In this study in the SN-erbB proteins we used the species as determined by SDS-PAGE crosslinking as a guide, and limited the species to monomer, dimer and trimer. In addition, we postulate that an unlimited number of proteins in a micelle would not be physically possible. Experimentally, this would behave as an aggregated protein, and we approximated this as a micelle containing ten monomers. While we did not observe any aggregation in our sedimentation equilibrium experiments, we wished to calculate the statistical distribution over wide range of mole fraction values. Including this parameter in the Poisson model

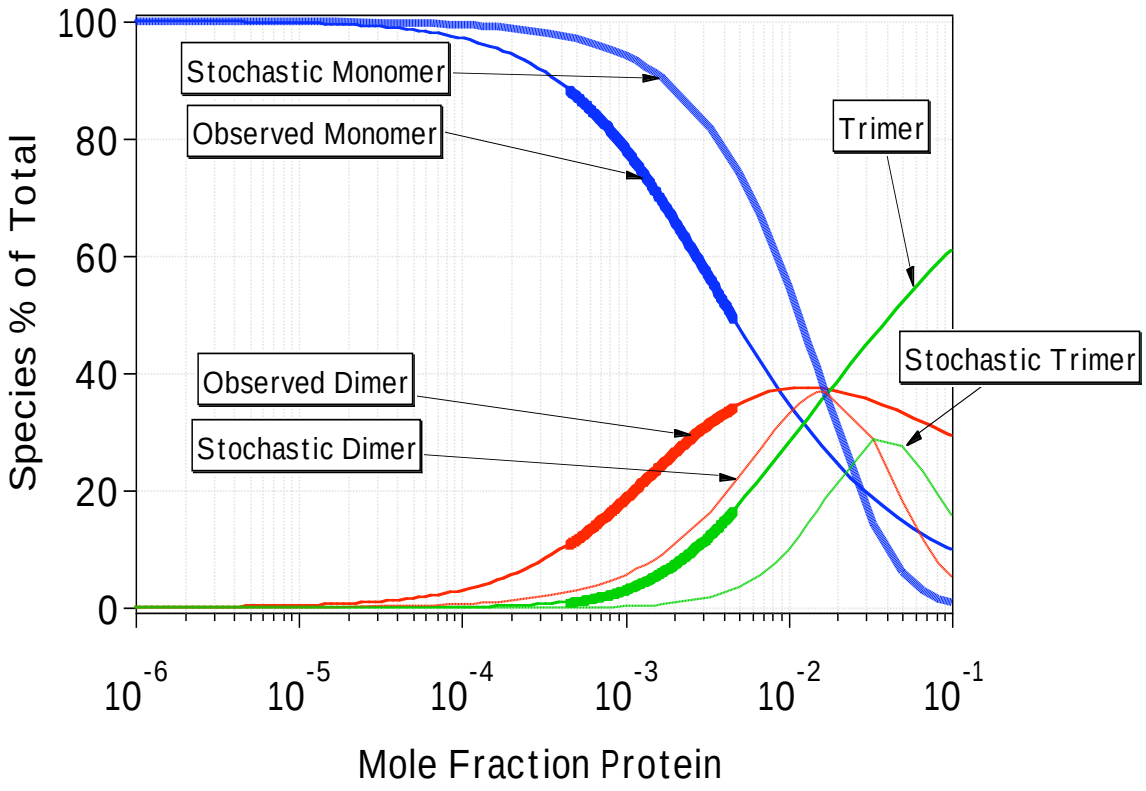


Figure 17. Species Distribution of SN-erbB4. It can be seen that the experimentally observed monomeric species of SN-erbB4 decays only slightly more rapidly that the random expectation, and that there is a significant non-negligible concentration of stochastic dimer and trimer under the experimental concentrations used. The species distributions for the SN-erbB1, SN-erbB2 and SN-erbB3 proteins are essentially the same as the random distributions shown here. Thus, we must conclude that the free energy of preferential interactions for these three transmembrane domains (as isolated domains) are extremely modest and on the order of 1 kcal mol^{-1} or less.

The blue, red and green lines represent monomeric, dimeric and trimeric species, respectively. The lines with a thickened portion represent the experimentally observed species distributions. The lines that are completely thin represent the stochastic distributions calculated from our theoretical model.

allows us to approximate when this might occur, and by including this term in the partition function, we confirmed that decamer is not a significant species within the concentration range shown in Figure 17 (data not shown). Thus, for clarity, we subsequently removed this term, and the final partition function contains only monomeric, dimeric and trimeric species.

Since we observed the most self-association for the SN-erbB4 construct, we found it most useful to compare the distribution of the species for this transmembrane domain with the expected random distribution. This is shown in Figure 17.

2.20.3. Sensitivity analysis of the micelle aggregation number in calculating the random distribution of oligomers.

A parameter in this statistical model required to convert micelles to total micellar detergent concentration is the micelle aggregation number, where the latter is the experimental variable. Because of questions regarding the impact of this number on the random distribution, we used a partition function containing only monomeric and dimeric species to determine the dependence of the random distribution on the aggregation number. To compare the results of our simulation with a simple monomer-dimer experimental system, we explored this issue in depth on a study of the interactions mediated by the transmembrane domain of the CCK-4 oncogene. CCK-4 is a recently cloned colon carcinoma kinase and it is thought to represent a new member of the receptor tyrosine kinase super family [21]. Like the transmembrane domains of the erbB receptors, the transmembrane domain of CCK-4 contains a conserved GxxxG sequence thought to promote interactions between transmembrane helices [22, 23]. The evolutionary conservation of the CCK-4 transmembrane sequence suggested the hypothesis that it may play a functionally important role in signaling that involves transmembrane helix-helix interactions. Like the SN-erbB receptors, we found extremely modest interactions between the CCK-4 transmembrane domains.

To calculate the dependence of the SN-CCK-4 random distribution on micelle aggregation number, we assumed a range of aggregation values. For protein-free micelles, we assumed an aggregation number for C14 betaine micelles of 100, which lies intermediate between the published values for C12 betaine and C16 betaine micelles [24, 25]. For protein-containing micelles, we recognize that the number of detergents bound may or may not be similar to the protein-free aggregation number, and it may also be that the monomeric and dimeric forms of the protein would be solvated by a micelle of a different size. Under the density-matching conditions used for sedimentation equilibrium, there is no experimental information about the size of the micelle required to solvate the transmembrane regions of the protein molecules. For large membrane proteins le Maire and colleagues have postulated that the amount of detergent bound depends on the hydrophobic protein surface area [26-28]. However, for smaller membrane proteins, such as the single transmembrane helix of CCK-4, it is thought that the transmembrane region may serve as an organizing hydrophobic nucleus surrounded by detergent molecules that bind in a micellar fashion, and the amount of detergent bound to the glycophorin A transmembrane helix is similar to the aggregation number for protein-free micelles [29]. We therefore used the protein-free number as a starting approximation of the number of detergent molecules bound, and

then we varied this number from 50-150 to account for our uncertainty. As long as the number of detergents bound is within this range, our results shown in Figure 17 indicate that the exact value is not an extremely sensitive factor in this analysis, and the fundamental conclusion is the same.

Figure 18 shows the mole fraction distribution of species for SN-CCK-4 overlaid upon the mole fraction distribution of species from the random distribution model containing monomer and dimer forms. It can be seen that the decrease in

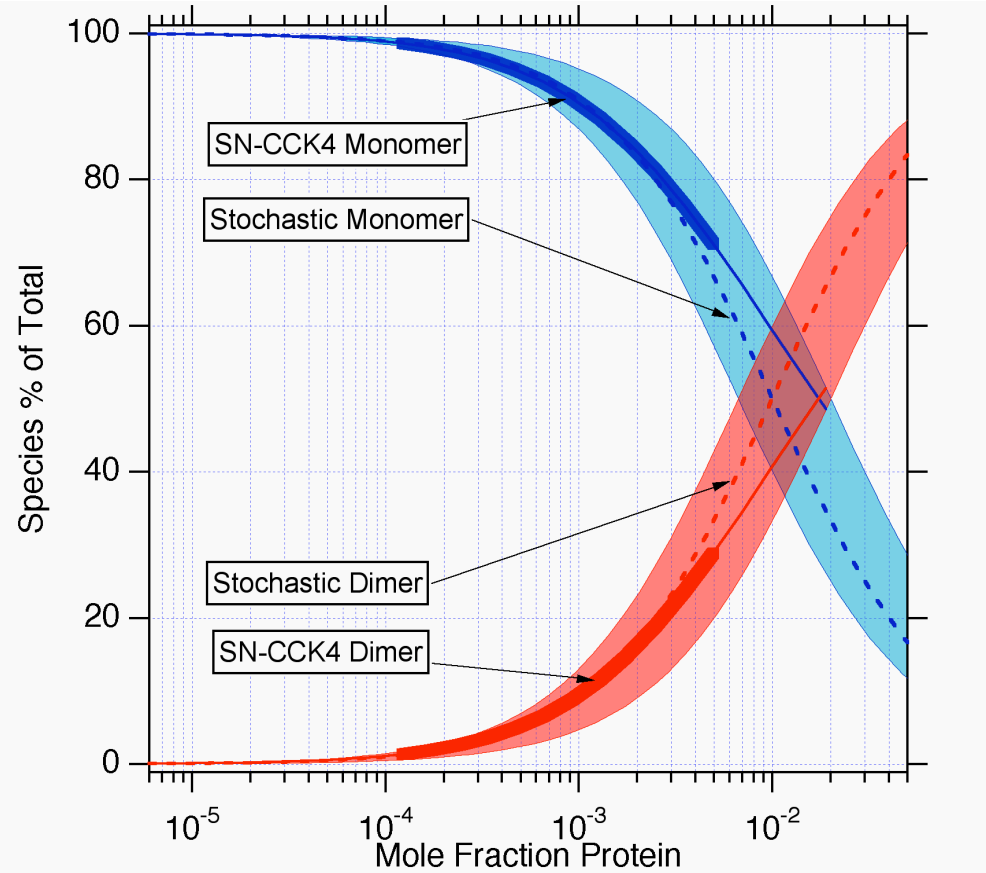


Figure 18. The SN-CCK4s species distribution is similar to the stochastic distribution in micelles.

The solid blue and red lines are the experimentally determined monomer and dimer distributions, respectively. The thickened portions of these curves represent those concentrations that were experimentally observed, and the thin portions of these curves represent extrapolations of those distributions to the equilibrium midpoint. The broken blue and red lines are the monomer and dimer distributions predicted from the Poisson distribution of proteins in micelles assuming no preferential interaction with an aggregation number of 100. The shaded regions represent the area covered by the distributions allowing the aggregation number to vary from 50-150. Mole fraction on the abscissa is defined as the ratio of moles of protein to moles of micellar detergent.

experimentally observed monomeric SN-CCK4s tracks well with the decrease in the statistical distribution of monomer using an aggregation number of 100 as depicted by the blue broken line. Concomitantly, the increase in experimentally observed dimeric SN-CCK4 is well approximated by the expected statistical distribution of dimer. Even if the number of detergents bound varies by ±50, the experimentally observed SN-CCK4s species deviate only slightly from the random distributions suggesting that the number of detergent molecules solvating each species is a relatively insensitive parameter. The results are very similar when the partition function for the random model includes an additional trimeric form (data not shown).

We can consider the free energy of preferential interactions for SN-CCK4 as a reflection of the deviation of the experimental monomer-dimer midpoint from that of the stochastic distribution. Within the range of distributions shown by the shaded regions of Figure 18, the comparison demonstrates that the free energy of any preferential self-interactions between SN-CCK4 proteins in micelles is minimal and certainly less than one kcal mol^{-1}. By comparison with the calculated random distributions, it can be seen that the midpoints overlay almost exactly if the average aggregation number equals 50. In this case, we can conclude that the dimerization propensity of SN-CCK4s arises solely from the random distribution in micelles; *e.g.* association of SN-CCK4 is completely due to overloading of the micelles with protein and not from any preferential interactions between the helices. On the other hand, if the average aggregation number is 150, the midpoint of the experimental curve ($1.5x10^{-2}$) falls to right of the stochastic curve ($5x10^{-3}$). This result might suggest the presence of slight repulsive interactions between the SN-CCK4 proteins. Physically, this is possible using this stochastic model, because there is a significant population of protein-free micelles at the midpoint of the random distribution. In this latter case, the mole fraction concentration difference between the midpoints of the random and experimental populations is approximately three-fold. From this, we can estimate the free energy of repulsion as being on the order of the available thermal energy at 25˚C. Thus, even if with the uncertainty of knowing the size of the protein-micelle complex, we can conclude that the transmembrane domains of CCK4, like those of the SN-erbB receptors and unlike those of glycophorin A, do not encode a thermodynamically meaningful preferential self-interaction in micelles.

2.20.4. A transmembrane domain lacking preferential interactions can still enhance the interactions of a covalently linked soluble domain: Reinterpretation of the Tanner & Kyte data and Implications for the affinity of EGF for the EGF receptor ligand-binding domain.

In studies of SNGpA fusion proteins, it has been shown that it is appropriate to treat the micelles as a distinct thermodynamic phase and to consider the transmembrane segments confined to this phase [9, 11, 30]. Such a consideration leads to the expression [11]:

$$\Delta G_{app} = \Delta G_x^o + RT \ln\left[MicellarDet\right]_w \qquad (6)$$

where ΔG_{app} is the observed apparent free energy in bulk aqueous concentrations, ΔG_x^o is the standard state mole fraction free energy of association, R is the universal gas constant, T is the temperature (K), and *[Micellar Det]$_w$* is the aqueous concentration of the detergent in micelles. The first term represents the association free energy of the intrinsic chemical components of the association reaction, and the second term represents a statistical energy of mixing, the magnitude of which varies with the micellar detergent concentration. In the case of GpA, we believe the contribution of the soluble SN domain to the standard state free energy, ΔG_x^o is negligible, and we interpret ΔG_x^o as the contribution from interactions between the transmembrane domains.

However, the soluble domain of EGFR (erbB1), when ligand-bound, **does interact**. Equation [5] predicts that extending these fragments to include a hydrophobic transmembrane segment should enhance the apparent dimerization affinity of the ECDs *even in the absence of specific TM interactions*. The TM domain would be restricted to the volume of the micellar phase, and thus the apparent free energy of association of the ECDs would now contain a contribution from the statistical energy term described above as well as the intrinsic free energy of dimerization of the soluble domains, ΔG_{int}. Applying the above formalism, we can substitute ΔG_{int}, for ΔG_x^o and rearranging equation [5] to write the apparent association equilibrium constant for an ECD-TM construct in micelles, K_{app}, as:

$$K_{app} = \frac{K_{int}}{[micellar\ \ Det]_w} \qquad (7)$$

The equilibrium constants in equation [6] are in bulk molar concentration units. From eqn [6], one would expect that the association of the receptor would depend on detergent concentration, with a tighter apparent association expected at lower detergent concentrations, and it has in fact been shown that the full-length EGFR, as well as a truncated fragment resembling v-erbB, interact more strongly at low detergent concentrations [31, 32].

A study by Tanner & Kyte, which compared the ligand-induced dimerization of the ECD of erbB1 to the dimerization of a fragment extended to include the TM, is also consistent with this formalism. This study found the longer fragment did indeed associate more readily, although complete dimerization was never observed [33]. The dimerization of the longer fragment was measured in 1% Triton X-100 (15mM), and equation [6] predicts that restriction of the soluble fragment to the micellar phase would lead to an apparent 67-fold enhancement in the association. By comparing the dimerization of the ECD alone, which had been shown to be complete at 40μM protein [34], with the dimerization of the ECD-TM fragment in Triton-micelles that they observed (2nM), the authors estimated an observed enhancement of approximately 10,000-fold [33]. However, a subsequent study reports complete ligand-induced dimerization of the extracellular domain of erbB1 at concentrations as low as 4μM protein [18], so the enhancement of ligand-induced dimerization of the ECD by the addition of the TM may

be less than originally suggested and be more on the order of 1,000-fold. Equation [6] suggests a large portion of this is simply the result of restricting the longer fragment to be stably inserted in the micelles. If approximately 65-fold arises from limiting the protein to the micellar phase, only a 15-30 fold enhancement arises from other features of the ECD-TM construct. This additional enhancement is modest, and it could arise from stabilization in the longer construct of structures in the juxtamembrane region important for dimerization or from restriction of the relative orientation of two ECDs on the surface of a micelle. The role of the transmembrane domain in the differences observed in the Tanner and Kyte study is still unclear, but our results suggest that it may not be necessary to invoke strong specific interactions between the TM domains to account for the observed enhancement.

2.21. Development of protocols for evaluating transmembrane protein-protein interactions in lipid bilayers (Task 2, Part C)

In light of the modest interactions we observed for the erbB transmembrane domain interactions in detergent micelles, we find it compelling to develop experimental approaches for the analysis of membrane protein interactions in lipid bilayers. The experimental approach will take advantage of the energy transfer that occurs when a donor fluorescent molecule comes within proximity of an acceptor molecule. We have been trying to use a control protein, the outer membrane dimerizing membrane protein OMPLA, as a control for development of this fluorescence assay, however we have not yet observed any energy transfer in any experiments.

2.22. Abbreviations Used:

SN:	Staphylococcal nuclease
GpA:	glycophorin A
SDS-PAGE:	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
IPTG	Isopropyl-D–thiogalactopyranoside
TM	Transmembrane

3. Key Research Accomplishments

- Cloning, expression and purification of all four of the TM domains of the human erbB receptors. (Task 1)
- Demonstration of weak self-association interactions for all four erbB TM domains in detergent micelles. (Task 2)
- Interpretation of Self-association energetics using a theoretical model that takes into account the stochastic interactions of TM domains in micelles (Task 3)

4. Reportable Outcomes

4.1. Clones

Thirteen clones were generated by these studies. These are made freely available to researchers who request them.

4.2. Three manuscripts were published in the granting period.

1. FJ Kobus & KG Fleming (2005) "The GxxxG-containing transmembrane domain of the CCK-4 oncogene does not encode preferential self-interactions" *Biochemistry* **44**: 1464-1470.
2. AM Stanley & KG Fleming (2005) "The transmembrane domains of erbB receptors do not dimerize strongly in micelles" *J. Mol. Biol.* **347**: 759-72.
3. KG Fleming (2005) "Analysis of membrane proteins using analytical ultracentrifugation" (Invited Book Chapter) in *Analytical Ultracentrifugation: Techniques and Methods* DJ Scott, SE Harding and AJ Rowe, Royal Society Publishing, pp 432-446.

4.3. Sixteen Oral Presentations of this work at Invited Seminars and National Meetings

2006

1. *Invited Speaker*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* Gordon Research Conference on Reversible Associations in Structural and Molecular Biology, Ventura, CA, January 2006.

2. *Scheduled Seminar*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* Virginia Commonwealth University, Department of Biochemistry, Richmond, VA, March 2006.

3. *Scheduled Seminar*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* University of Maryland, Department of Chemistry and Biochemistry, College Park, MD, April 2006.

4. *Invited Speaker*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* Gordon Research Conference on Biopolymers, RI, June, 2006.

5. *Invited Speaker*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* FASEB Summer Research Conference on Membrane Molecular Biophysics, Saxtons River, VT, July, 2006.

6. *Scheduled Seminar*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* University of Glasgow, Glasgow, UK, April 2006

7. *Scheduled Seminar*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* University of Oxford, Division of Structural Bioinformatics and Computational Biochemistry, Oxford, UK, April 2006

8. *Scheduled Seminar*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* University of York, Division of Biochemistry, York, UK, April 2006

2005

7. *Speaker,* (Karen Fleming) *"Analysis of Membrane protein interactions using the analytical ultracentriuge"* The Ninth Symposium on Modern Analytical Ultracentrifugation, University of Connecticut, June 17, 2005.

8. *Keynote Speaker*, Howard Hughes Symposium, University of Richmond, Richmond, VA, September 2005.

9. *Scheduled Seminar*, Department of Chemistry, City University of New York, November 2005.

2004

10. *Scheduled Seminar,* (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* Kansas University Medical Center, October 2004.

11. *Scheduled Seminar*, (Karen Fleming) *"Thermodynamics of Membrane protein interactions"* Department of Biochemistry and Cell Biology, SUNY Stony Brook, NY, November 2004.

2003

12. Karen G. Fleming, Speaker (2003) "Membrane Protein Interactions" The Seventeenth Symposium of the Protein Society, Boston, MA, July 2003.
13. Karen G. Fleming*, Scheduled Seminar*, Laboratory of Chemical Physics, National Institutes of Health, Bethesda, MD, July 23, 2003.
14. Karen G. Fleming*, Scheduled Seminar*, University of Vermont, Department of Biochemistry, Burlington, VT, January 17, 2003.
15. Karen G. Fleming*, Scheduled Seminar*, Department of Biological Chemistry, Johns Hopkins Medical School, May 6, 2003.

16 Karen G. Fleming*, Scheduled Seminar*, Carnegie Institute of Washington, Baltimore, MD, September 22, 2003.

4.4. One Award for this Research

2004

Ann Marie Stanley*, Maridel Lares, and Karen G. Fleming (2004) "Interactions of the Transmembrane Domains of the EGF Receptors" UK International Analytical Ultracentrifuge Meeting, Oxford, UK, March 2004.

**Ann Marie Stanley, the first author for this poster, won the Best Poster award for this work at the UK meeting.*

5. List of personnel receiving pay from this award

Karen G. Fleming, PI
Ann Marie Stanley, Graduate student
Maridel Lares, Research Assistant

6. Conclusions.

Our results represent the first quantitative measurement of the self-association propensity for these erbB receptor transmembrane domains. All erbB transmembrane domains drive weak self-association of the SN-chimeric construct in 11 mM C_8E_5 micelles. If the concentration of detergent micelles is increased, the proteins are effectively diluted. In 23 mM C_8E_5, the SNerbB1, SNerbB2 and SNerbB3 proteins are monomeric whereas the SNerbB4 protein still shows some level of association under this condition. Upon increasing the detergent concentration further to 50 mM C_8E_5, the SNerbB4 protein is completely monomeric. We developed a theoretical model to explain the expected random distribution of oligomers in micelles, and we find that the SN-erbB1, SN-erbB2 and SN-erbB3 species distributions are well described by this model whereas the SN-erbB4 distribution shows a slight amount of thermodynamically

meaningful preferential association. We estimate the free energy of any preferential interactions to be on the order of 1 kcal mol^{-1} or less.

We applied our thermodynamic model to analyze the previously published data of Tanner & Kyte [33]. We find that the covalent attachment of a transmembrane domain to a soluble (ligand-binding) domain can enhance the interactions of the soluble domain. Thus, strong preferential interactions between transmembrane domains are not required. We propose that interactions between the transmembrane domains need only be sterically permissible, and that the full-length receptor may have many sites of weak interactions.

7. References:

1. Riese, D.J., 2nd and D.F. Stern, *Specificity within the EGF family/ErbB receptor family signaling network.* Bioessays, 1998. **20**(1): p. 41-8.
2. Slamon, D.J., et al., *Studies of the HER-2/neu proto-oncogene in human breast cancer and ovarian cancer.* Science, 1989. **244**(707-712).
3. Muss, H.B., *c-erbB-2 expression and resonse to adjuvant therapy in women with node-positive early breast cancer.* N. Engl. J. Med., 1994. **330**: p. 1260-1266.
4. Fleming, K.G. and D.M. Engelman, *Specificity in transmembrane helix-helix interactions defines a hierarchy of stability for sequence variants.* Proc. Natl. Acad. Sci. USA, 2001. **98**: p. 14340-14344.
5. Bargmann, C.I. and R.A. Weinberg, *Increased Tyrosine Kinase Activity Associated with the Protein Encoded by the Activated neu Oncogene.* Proc. Natl. Acad. Sci. USA, 1988. **85**: p. 5394-5398.
6. Bargmann, C.I. and R.A. Weinberg, *Oncogenic Activation of the neu-encoded Receptor Protein by Point Mutation and Deletion.* EMBO J., 1988. **7**: p. 2043-2052.
7. Mendrola, J.M., et al., *The single transmembrane domains of ErbB receptors self-associate in cell membranes.* J Biol Chem, 2002. **277**(7): p. 4704-12.
8. Bargmann, C., M. Hung, and R. Weinberg, *Multiple Independent Activations of the neu Oncogene by a Point Mutation Altering the Transmembrane Domain of p185.* Cell, 1986. **45**: p. 649-657.
9. Doura, A.K. and K.G. Fleming, *Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer.* J Mol Biol, 2004. **343**: p. 1498-1497.
10. Doura, A.K., et al., *Sequence context modulates the stability of a GxxxG mediated transmembrane helix-helix dimer.* J Mol Biol, 2004. **341**: p. 991-998.
11. Fleming, K.G., *Standardizing the free energy change of transmembrane helix-helix interactions.* J Mol Biol, 2002. **323**: p. 563-571.
12. Fleming, K.G., et al., *Thermodynamics of glycophorin A transmembrane helix-helix association in C14 betaine micelles.* Biophys Chem, 2004. **108**: p. 43-49.
13. Fleming, K.G., A.L. Ackerman, and D.M. Engelman, *The Effect of Point Mutations on the Free Energy of Transmembrane α-Helix Dimerization.* J. Mol. Biol., 1997. **272**: p. 266-275.
14. Fleming, K.G., *Measuring transmembrane α-helix energies using analytical ultracentrifugation*, in *Chemtracts: biological applications of the analytical*

ultracentrifuge, J.C. Hanson, Editor. 1998, Springer-Verlag: New York. p. 985-990.
15. Yphantis, D.A., *Equilibrium Ultracentrifugation of Dilute Solutions.* Biochemistry, 1964. **3**: p. 297-317.
16. Tzahar, E., et al., *A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor.* Mol Cell Biol, 1996. **16**(10): p. 5276-87.
17. Pinkas-Kramarski, R., et al., *Diversification of Neu differentiation factor and epidermal growth factor signaling by combinatorial receptor interactions.* Embo J, 1996. **15**(10): p. 2452-67.
18. Ferguson, K.M., et al., *Extracellular domains drive homo- but not hetero-dimerization of erbB receptors.* Embo J, 2000. **19**(17): p. 4632-43.
19. Tanford, C. and J.A. Reynolds, *Characterization of Membrane Proteins in Detergent Solutions.* Biochim. Biophys. Acta, 1976. **457**: p. 133-170.
20. Hall, D. and A.P. Minton, *Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges.* Biochim Biophys Acta, 2003. **1649**(2): p. 127-39.
21. Mossie, K., et al., *Colon carcinoma kinase-4 defines a new subclass of the receptor tyrosine kinase family.* Oncogene, 1995. **11**(10): p. 2179-84.
22. Senes, A., M. Gerstein, and D.M. Engelman, *Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions.* J Mol Biol, 2000. **296**(3): p. 921-36.
23. Russ, W.P. and D.M. Engelman, *The GxxxG motif: a framework for transmembrane helix-helix association.* J Mol Biol, 2000. **296**(3): p. 911-9.
24. Lichtenberg, D., R.J. Robson, and E.A. Dennis, *Solubilization of phospholipids by detergents. Structural and kinetic aspects.* Biochim Biophys Acta, 1983. **737**(2): p. 285-304.
25. Neugebauer, J., *A guide to the properties and uses of detergents in biology and biochemistry*. 1990, San Diego: Calbiochem Corporation.
26. Møller, J.V. and M. le Maire, *Detergent binding as a measure of hydrophobic surface area of integral membrane proteins.* J Biol Chem, 1993. **268**: p. 18659-18672.
27. le Maire, M., P. Champeil, and J.V. Møller, *Interaction of membrane proteins and lipids with solubilizing detergents.* Biochim Biophys Acta, 2000. **1508**: p. 86-111.
28. le Maire, M., et al., *Mode of interaction of polyoxyethyleneglycol detergents with membrane proteins.* Eur J Biochem, 1983. **129**: p. 525-532.
29. Grefrath, S.P. and J.A. Reynolds, *The molecular weight of the major glycoprotein from the human erythrocyte membrane.* Proc. Natl. Acad. Sci. USA, 1974. **71**(10): p. 3913-3916.
30. Lomize, A.L., I.D. Pogozheva, and H.I. Mosberg, *Quantification of helix-helix binding affinities in micelles and lipid bilayers.* Protein Sci, 2004. **13**(10): p. 2600-12.
31. Kwatra, M.M., D.D. Bigner, and J.A. Cohn, *The ligand binding domain of the epidermal growth factor receptor is not required for receptor dimerization.* Biochim Biophys Acta, 1992. **1134**(2): p. 178-81.

32. Moriki, T., H. Maruyama, and I.N. Maruyama, *Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain.* J Mol Biol, 2001. **311**(5): p. 1011-26.
33. Tanner, K.G. and J. Kyte, *Dimerization of the extracellular domain of the receptor for epidermal growth factor containing the membrane-spanning segment in response to treatment with epidermal growth factor.* J Biol Chem, 1999. **274**(50): p. 35985-90.
34. Burke, C.L., et al., *Dimerization of the p185neu transmembrane domain is necessary but not sufficient for transformation.* Oncogene, 1997. **14**: p. 687-696.

8. Appendices:

1. FJ Kobus & KG Fleming (2005) "The GxxxG-containing transmembrane domain of the CCK-4 oncogene does not encode preferential self-interactions" *Biochemistry* **44**: 1464-1470.

2. AM Stanley & KG Fleming (2005) "The transmembrane domains of erbB receptors do not dimerize strongly in micelles" *J. Mol. Biol.* **347**: 759-72.

3. Proofs for the following book chapter. No electronic copy of the book chapter in the published form is available. A Xerox copy can be sent by regular mail if desired.
KG Fleming (2005) "Analysis of membrane proteins using analytical ultracentrifugation" (Invited Book Chapter) in *Analytical Ultracentrifugation: Techniques and Methods* DJ Scott, SE Harding and AJ Rowe, Royal Society Publishing, pp 432-446.

The GxxxG-Containing Transmembrane Domain of the CCK4 Oncogene Does Not Encode Preferential Self-Interactions†

Felix J. Kobus and Karen G. Fleming*

T. C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218

Received September 7, 2004; Revised Manuscript Received November 21, 2004

ABSTRACT: The recently cloned colon carcinoma kinase 4 (CCK4) oncogene contains an evolutionarily conserved GxxxG motif in its single transmembrane domain (TMD). It has previously been suggested that this pairwise glycine motif may provide a strong driving force for transmembrane helix−helix interactions. Since CCK4 is thought to represent a new member of the receptor tyrosine kinase family, interactions between the TMDs may be important in receptor self-association and activation of signal transduction pathways. To determine whether this conserved CCK4 TMD can drive protein−protein interactions, we have carried out a thermodynamic study using the TMD expressed as a *Staphylococcal* nuclease (SN) fusion protein. Similar SN−TMD fusion proteins have been used to determine the sequence specificity and thermodynamics of transmembrane helix−helix interactions in a number of membrane proteins, including glycophorin A. Using sedimentation equilibrium in C14 betaine micelles, we discovered that the CCK4 TMD is unable to drive strong protein−protein interactions. At high protein/detergent ratios, the SN−CCK4 fusion protein will dimerize, but a stochastic model for protein association in micelles can explain the observed dimer population. For low-affinity interactions such as the one studied here, an understanding of this discrete stochastic distribution of membrane proteins in micelles is important for distinguishing between preferential and random self-interactions, which can both influence the oligomeric population. The lack of a thermodynamically meaningful self-association propensity for the CCK4 TMDs demonstrates that a GxxxG motif is not sufficient to drive transmembrane helix−helix interactions.

Receptor tyrosine kinases (RTKs)[1] are known to play an important role in the development and/or progression of many forms of cancer. A molecular understanding of the mechanisms by which these receptors promote cellular transformation is critical for the understanding and treatment of human cancer. In normal cells, RTKs are transmembrane signaling proteins that transmit biological signals from the extracellular environment to the interior of the cell. The regulated transmission of these signals is important for many cellular events, such as cell growth and differentiation, axonal growth, epithelial growth and development (*1*). RTKs are misregulated in many human cancers including breast cancer (erbB2/Her2) (*2*, *3*), ovarian cancer (erbB2) (*2*, *3*), melanoma (CCK4) (*4*), and colon cancer (erbB1 and CCK4) (*5*, *6*). The predominant RTK-related genetic alteration in human cancer is not mutation of the receptor protein, but rather it is amplification of the gene. Subsequent overexpression of receptor proteins leads to constitutive stimulation of the RTK activity and uncontrolled cellular signaling.

Recently, colon carcinoma kinase 4 (CCK4), a new member of the receptor tyrosine kinase superfamily was cloned from human colon carcinoma-derived cell lines (*7*). The CCK4 mRNA for this gene was variably expressed in colon carcinoma derived cell lines but not expressed in human adult colon tissues. In contrast, it is expressed in fetal colon of the mouse (*7*). These observations suggest that the normal role for CCK4 may be in development of the colon followed by down regulation in the adult tissue. The unusual expression of CCK4 in colon carcinoma cell lines might indicate that the protein is functioning abnormally in adult colon cells, which suggests a role in colon carcinoma development and/or proliferation.

Orthologues for CCK4 have been cloned from other species (*8*−*10*). Amino acid residues within the transmembrane domain (TMD) are among the conserved features (*9*). The CCK4 TMD is 55% identical to the transmembrane sequences of the chicken *Klg* and Hydra *Lemon* orthologues (*9*), and a pattern of residues with helical periodicity is conserved. Within the conserved residues of the TMD is a GxxxG sequence, a motif thought to promote interactions between transmembrane helices (*11*, *12*). The evolutionary conservation of the CCK4 transmembrane sequence suggests the hypothesis that it may play a functionally important role in signaling that involves transmembrane helix−helix interactions. To test whether this CCK4 TMD may serve as a driving force for protein−protein interactions, we carried out a thermodynamic study of the TMD using sedimentation equilibrium in detergent micelles.

† This work was supported by a grant from the NIH (GM57534) and by a Career Award from the Department of Defense (DAMD17-02-1-0427).

* To whom correspondence should be addressed. Voice: 410-516-7256. Fax: 410-516-4118. E-mail: Karen.Fleming@jhu.edu

[1] Abbreviations: CCK4, colon carcinoma kinase 4; C14 betaine, 3-(*N*,*N*-dimethylmyristyl-ammonio)propanesulfonate; RTK, receptor tyrosine kinases; SN, *Staphylococcal* nuclease; TMD, transmembrane domain; GpA, glycophorin A; SN−CCK4, a fusion protein composed of SN and the transmembrane domain of CCK4; SN−GpA, a fusion protein composed of SN and the transmembrane domain of glycophorin A.

10.1021/bi048076l CCC: $30.25
Published on Web 01/08/2005

Table 1: Transmembrane Sequences Used in This Study[a]

Protein	SN-TM fusion sequence aligned by the GxxxG motif
CCK4s	QLAVHHFSEPG IGLSVGAAVAYIIAVLGLMFYSKKR
CCK4xs	QLAVHHFSEPGMIQTIGLSVGAAVAYIIAVLGLMFYSKKR
	* * ** * *** * *** **
GpA WT	QLAVHHFSEPEITLIIFGVMAGVIGTILLISYGIRRLI

[a] Sequences are aligned by the GxxxG motif. The predicted transmembrane domains are underlined. The N-terminal sequence shows the junction between the SN sequence and the cloned transmembrane domain of interest. The conserved residues in the transmembrane domain are denoted with a star. In a manner analogous to the GpA construct, positive charges were added to the C-termini of the CCK4 sequences. The SN−CCK4xs contains the extended linker region.

EXPERIMENTAL METHODS

Cloning and Sample Preparation. An image clone for the CCK4 protein was purchased from Invitrogen. The sequence encoding the transmembrane region was amplified using the polymerase chain reaction using primers containing unique SmaI and BamH I sites, and the PCR product was cloned C-terminal to *Staphylococcal* nuclease (SN) in an IPTG-inducible pET11a expression construct from which the open reading frame for the glycophorin A transmembrane had been excised. The DNA sequence was verified by double-stranded DNA sequencing. The Cys-to-Ser mutation (immediately before the KKR) was introduced using the QuickChange mutagenesis kit. The SN−CCK4s and SN−CCK4xs proteins were purified using the published protocol for the SN−GpA proteins (*13*). Immediately before sedimentation equilibrium analysis, samples were exchanged by ion-exchange chromatography into buffer containing 10 mM C14 betaine. SigmaUltra grade C14 betaine was purchased from Sigma.

Sedimentation Equilibrium Analytical Ultracentrifugation. Sedimentation equilibrium experiments were performed at 25 °C using a Beckman XL-A analytical ultracentrifuge using six-sector cells equipped with quartz windows in an AnTi 60 rotor as described previously in detail (*13*−*15*). The samples were centrifuged for lengths of time sufficient to achieve equilibrium determined by using the MATCH algorithm. Equilibrium data obtained from absorbance at 280 nm were analyzed by nonlinear least-squares curve fitting of radial concentration profiles with the Macintosh version of NONLIN (*16*) using the equations describing the reversible association in sedimentation equilibrium. For each global fit, nine data sets were used consisting of three different initial protein concentrations (initial A_{280} absorbance values of 0.9, 0.6, and 0.3 cm^{-1}) analyzed at three significantly different speeds (20000, 24500, 30000 revolutions per minute). The monomeric molecular masses and partial specific volumes were calculated using the program SEDNTERP (*17*), and these parameters were held constant in fitting the absorbance versus radius profiles. The equilibrium constant was converted from absorbance to (bulk molar) units using molar extinction coefficients of 20952 and 21425 cm^{-1} M^{-1} for the SN−CCK4s and SN−CCK4xs proteins, respectively. These values were calculated by adjusting the experimental molar extinction coefficient for SN for the additional mass of the TMD of interest as has been previously done for the glycophorin A transmembrane domain (*18*).

The apparent free energy of association was calculated from the apparent monomer−dimer equilibrium constant in the standard way, $\Delta G^{\circ}_{App} = -RT \ln K_{App}$. The standard state mole fraction free energy of association was calculated using the ideal-dilute assumption as $\Delta G^{\circ}_{x} = -RT \ln K_x$ where $K_x = K_{App}[\text{micellarC14}]_w$ and $[\text{micellarC14}]_w$ is the concentration of micellar C14 betaine expressed on the aqueous molar scale (*14*, *15*).

Sequence Motifs. Sequence motifs were identified using the online TMSTAT calculation (*11*) (http://bioinfo.mbb.yale.edu/tmstat/) with the extended version of the CCK4 TMD.

RESULTS

We expressed the CCK4 TMD as a C-terminal fusion protein with *Staphylococcal* nuclease (SN), an experimental approach that has been successfully used to map the helix−helix interactions of glycophorin A (*13*−*15*, *19*−*22*), BNIP-3 (*23*), and phospholamban (*24*) TMDs. We have previously shown that this method using this type of SN fusion construct has resulted in a quantitative ranking of glycophorin A transmembrane mutants by their ability to dimerize (*13*, *21*, *22*), and the relative stabilities of a series of GpA alanine mutants is conserved between the micellar environment in vitro and bacterial membranes in vivo (*13*). The advantage of using the sedimentation equilibrium method is that the molecular mass and thus stoichiometries are determined. In addition, in the absence of protein−protein association, sedimentation equilibrium will still result in a positive, interpretable result since it will directly report the monomeric molecular weights.

The CCK4 GxxxG-Containing Transmembrane Domain Self-Associates Only Weakly. Table 1 shows the CCK4 transmembrane sequences considered in this study. Since the GxxxG motif in CCK4 is located closer to the predicted N-terminus of its TMD, we were concerned that self-association might be sterically hindered by the SN portion of the fusion protein. Therefore, we extended the linker region by adding four additional residues from the CCK4 open reading frame. In addition, the C-terminal end of the wild-type CCK4 TMD contains a cysteine residue, which we mutated to serine. For comparison, the sequence of the glycophorin A transmembrane WT TMDs is also shown. We first used SDS−PAGE to screen the purified proteins for the ability of the CCK4 TMD to drive strong self-association.

This method demonstrated that both of the SN−CCK4 fusion proteins migrate as a monomeric species on SDS−PAGE. In contrast, the SN−GpA fusion protein migrates predominantly as a dimer under those same conditions (data not shown).

To further test for a self-association propensity of the CCK4 TMD, we carried out sedimentation equilibrium analytical ultracentrifugation experiments. Sedimentation equilibrium allows detection of protein−protein interactions over a range of association strengths. In particular, we have detected significant dimer populations of the SN-GpA mutant proteins even when these proteins are monomeric as measured by SDS−PAGE (*13*, *18*, *21*, *22*). In addition, unlike SDS−PAGE, sedimentation equilibrium experiments can provide a direct measure of the equilibrium constant for a self-association reaction. We chose to use C14 betaine micelles since this detergent micelle environment has been used for the analysis of several designed transmembrane helices (*25*) along with the fact that the glycophorin A TMD dimer shows in ideal-dilute behavior in these micelles (*15*).

Sedimentation equilibrium experiments carried out in 10 mM C14 betaine detergent micelles suggest that the SN−CCK4s protein can self-associate, albeit weakly. In four independent experiments, the single ideal species fit returned a molecular weight value that is approximately 20% greater than the calculated value for the monomeric protein, which suggests the formation of higher order species. To determine the molecular weights of these self-associated forms, the data were fit to association models including dimers, trimers, or tetramers, In three of four of the experiments, the monomer−dimer fit is the best description of the data. Figure 1 shows a typical species distribution for this monomer−dimer fit, where it can be seen that approximately 20% of the SN−CCK4s protein is dimeric. In the fourth independent experiment, a monomer−trimer fit was equally as good as a monomer−dimer fit; however, the dimerization constant for the monomer−dimer fit of these data was consistent with the other three data sets and we therefore used this value.

The average apparent free energy of association determined in 10 mM C14 betaine micelles at 25 °C, ΔG°_{App}, equals −5.2 (±0.1) kcal mol^{-1} ($N = 4$). Under the ideal-dilute assumption (*14*), this apparent free energy would correspond to a standard state mole fraction association free energy, ΔG°_{x}, equal to −2.40 (±0.1) kcal mol^{-1}. This value is considerably weaker than that of the SN−GpA TMD in either C14 betaine micelles, $\Delta G^{\circ}_{x} = -5.8$ kcal mol^{-1} (*15*) or C_8E_5 micelles, $\Delta G^{\circ}_{x} = -7.0$ kcal mol^{-1} (*14*). Figure 2 shows a comparison of the dimer distributions for SN−CCK4s and SN−GpA calculated from the standard state free energy values in C14 betaine micelles where it can be seen that this free energy difference leads to significant shifts in the concentration range over which the protein population is predominantly dimeric. At a total protein concentration of 10 μM, the SN−GpA fusion protein is >80% dimeric in 10 mM C14 betaine micelles. In contrast, the SN−CCK4s protein is ∼10% dimeric at this same overall protein concentration in 10 mM C14 betaine micelles.

Because of the observation that the self-interactions were so weak for this first SN−CCK4 construct, we also carried out sedimentation equilibrium experiments using the SN−CCK4xs fusion protein containing an extended linker region between SN and the TMD. The results for four independent

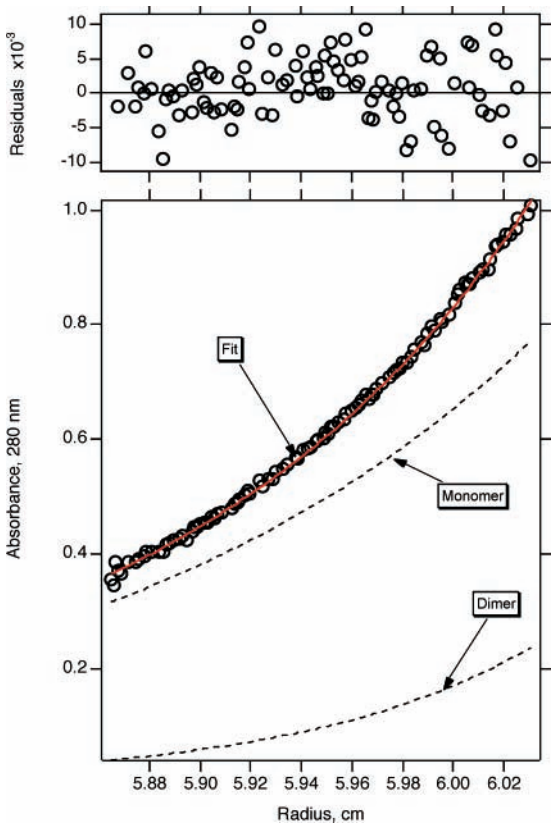


FIGURE 1: SN−CCK4s species distribution in C14 betaine micelles. The lower panel shows a representative data set for 43 μM (bulk) SN−CCK4s collected in 10 mM C14 betaine at 20000 rpm. The open circles are the data, and the line shows the global fit resulting from the global fit of nine data sets for this experiment. The broken lines represent the monomer and dimer contributions as determined from the global analysis, which sum to equal the fit. The upper panel shows the residuals of the fit for this data set, which are small and random.

sedimentation equilibrium experiments on the SN−CCK4xs protein were within error of the original SN−CCK4s fusion protein. The average apparent free energy of association, ΔG°_{App}, was equal to −5.3 (±0.2) kcal mol^{-1} ($N = 4$) in 10 mM C14 betaine micelles. Under the ideal-dilute assumption (*14*), this apparent free energy would correspond to a standard state mole fraction association free energy, ΔG°_{x}, equal to −2.50 (±0.2) kcal mol^{-1}. Thus, extending the linker region of the CCK4 TMD using native sequence resulted in no enhancement of protein−protein interactions.

DISCUSSION

To determine whether the GxxxG-containing TMD of CCK4 encodes the potential for strong helix−helix self-association, we carried out a thermodynamic study using an SN−CCK4 TMD fusion construct. This type of construct has been used in the past to map the sequence dependence of helix−helix oligomerization in several transmembrane proteins, most notably the TMD of glycophorin A (*13*, *19*−*21*). Using both SDS−PAGE and sedimentation equilibrium in detergent micelles, we find that the SN−CCK4 protein shows a very weak propensity for self-association.

The SN−CCK4 Species Distribution Is Well Described by a Random Distribution of Proteins in Micelles. To observe SN−CCK4 interactions, the current experiments were carried out at protein/detergent ratios much higher than were

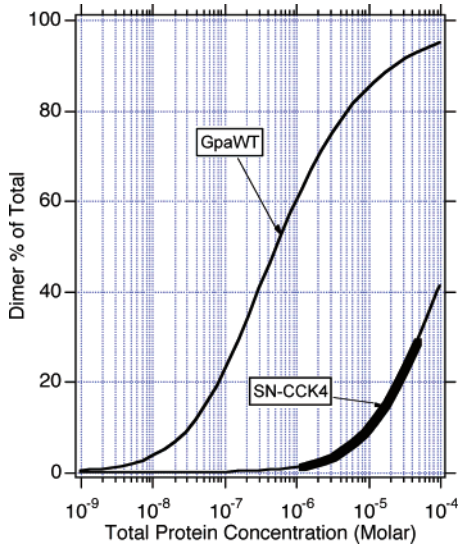


FIGURE 2: CCK4 is destabilized with respect to glycophorin A. A comparison of the dimer species distribution between the SN−GpA and SN−CCK4 transmembrane proteins in 10 mM C14 betaine micelles. The thin curves are calculated based on the free energies of association. The thickened portion of the SN−CCK4 curve represents the region of the data that was observed in the sedimentation equilibrium experiments. The less favorable free energy of association for SN−CCK4 shifts the population of dimer to higher protein concentrations.

previously used for glycophorin A (*15*), and at the highest SN−CCK4 protein concentration the number of protein molecules approached the number of micelles. Since the fundamental measurement in the sedimentation equilibrium under the current experimental conditions is the protein mass, two protein molecules that occupy the same micelle will sediment together and be measured as a dimer whether or not there is a preferential interaction between them. This co-partitioning phenomenon has been referred to in the literature as "artifactual togetherness" (*26*) and should only occur at high protein concentrations coupled with low detergent micelle concentrations. We hypothesize that this co-partitioning is due to an excluded volume effect and that these conditions can be thought of as thermodynamically nonideal. A similar excluded volume effect has been indicated to occur in soluble proteins at high protein concentrations or in the presence of small molecules that induce crowding conditions (*27*). To determine the extent of thermodynamically meaningful preferential interactions between the SN−CCK4 transmembrane proteins, it is first necessary to distinguish them from the protein interactions induced by the limiting solvent conditions.

To develop a model describing the distribution of non-interacting proteins in micelles, we employed a statistical approach that considers distribution of proteins in micelles in terms of the random occupancy of placing balls in bins, where the protein molecules represent balls and the individual micelles represent bins. Using this formalism, the random occupancy of any particular micelle by any particular protein molecule can be described as a Bernoulli event (e.g., either the micelle contains that protein or not), which can be described by a binomial distribution:

$$\Pr[X_i = k] = \binom{m}{k} p^k (1-p)^{m-k} \tag{1}$$

where X_i is the random variable that counts the number of proteins in a particular micelle, i, k is the number of proteins occupying a micelle (e.g., the "oligomeric state"), m is the number of protein molecules, p is the probability of each Bernouilli trial, which equals $1/n$, and n is the number of micelles. When the number of proteins and the number of micelles both exceed the allowed occupancy of a particular micelle, k, this binomial distribution is approximated by the Poisson distribution:

$$\lim (m,n \gg k):\Pr[X_i = k] \approx \frac{1}{k!}\left(\frac{m}{n}\right)^k e^{-m/n} \tag{2}$$

To consider the statistical protein distribution, we write the partition function, P, as the sum all those probabilities for which $k > 0$. (e.g., for this purpose, we are not interested in the distribution of "empty" micelles). To convert each probability to a relative concentration, each term is multiplied by the number of proteins that the probability represents.

$$P = \sum_{k>0} k\Pr[X_i = k] \tag{3}$$

The fraction of any particular species can then be calculated as a function of the mole fraction protein (m/n) in the micellar phase from this partition function by dividing its contribution by the sum of all species. For example, the fraction dimer can be calculated as

$$f_{\text{Dimer}} = \frac{2\Pr[X_i = 2]}{\sum_{k>0} k\Pr[X_i = k]} \tag{4}$$

The exact form of the partition function and the distribution of species will depend on the allowed values of k in the model. In this study, we limited the species to monomer and dimer. In addition, we postulate that an unlimited number of proteins in a micelle would not be physically possible. Experimentally, this would behave as an aggregated protein, and we approximated this as a micelle containing 10 monomers. While we did not observe any aggregation in our sedimentation equilibrium experiments, we wished to calculate the statistical distribution over wide range of mole fraction values. Including this parameter in the Poisson model allows us to approximate when this might occur, and by including this term in the partition function, we confirmed that decamer is not a significant species within the concentration range shown in Figure 3 (data not shown). Thus, for clarity, we subsequently removed this term, and the final partition function contains only monomeric and dimeric species.

A final parameter in this statistical model is required to convert micelles to total micellar detergent concentration, where the latter is the experimental variable. For protein-free micelles, we assumed an aggregation number for C14 betaine micelles of 100, which lies intermediate between the published values for C12 betaine and C16 betaine micelles (*28*, *29*). For protein-containing micelles, we recognize that the number of detergents bound may or may not be similar to the protein-free aggregation number, and it may also be that the monomeric and dimeric forms of the protein would be solvated by a micelle of a different size. Under the

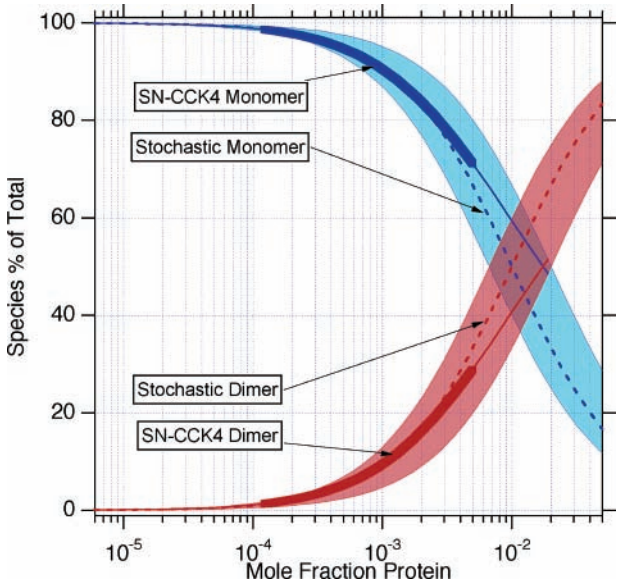


FIGURE 3: The SN−CCK4s species distribution is similar to the stochastic distribution in micelles. The solid blue and red lines are the experimentally determined monomer and dimer distributions, respectively. The thickened portions of these curves represent those concentrations that were experimentally observed, and the thin portions of these curves represent extrapolations of those distributions to the equilibrium midpoint. The broken blue and red lines are the monomer and dimer distributions predicted from the Poisson distribution of proteins in micelles assuming no preferential interaction with an aggregation number of 100. The shaded regions represent the area covered by the distributions allowing the aggregation number to vary from 50 to 150. Mole fraction on the abscissa is defined as the ratio of moles of protein to moles of micellar detergent.

density-matching conditions used for sedimentation equilibrium, there is no experimental information about the size of the micelle required to solvate the transmembrane regions of the protein molecules. For large membrane proteins le Maire and colleagues have postulated that the amount of detergent bound depends on the hydrophobic protein surface area (*30*−*32*). However, for smaller membrane proteins, such as the single transmembrane helix of CCK4, it is thought that the transmembrane region may serve as an organizing hydrophobic nucleus surrounded by detergent molecules that bind in a micellar fashion, and the amount of detergent bound to the glycophorin A transmembrane helix is similar to the aggregation number for protein-free micelles (*33*). We therefore used the protein-free number as a starting approximation of the number of detergent molecules bound, and then we varied this number from 50 to 150 to account for our uncertainty. As long as the number of detergents bound is within this range, our results shown in Figure 3 indicate that the exact value is not an extremely sensitive factor in this analysis, and the fundamental conclusion is the same.

Figure 3 shows the mole fraction distribution of species for SN−CCK4 overlaid upon the mole fraction distribution of species from the random distribution model containing monomer and dimer forms. It can be seen that the decrease in experimentally observed monomeric SN−CCK4s tracks well with the decrease in the statistical distribution of monomer using an aggregation number of 100 as depicted by the blue broken line. Concomitantly, the increase in experimentally observed dimeric SN−CCK4 is well approximated by the expected statistical distribution of dimer. Even if the number of detergents bound varies by ±50, the experimentally observed SN−CCK4s species deviate only slightly from the random distributions, suggesting that the number of detergent molecules solvating each species is a relatively insensitive parameter. The results are very similar when the partition function for the random model includes an additional trimeric form (data not shown).

We can consider the free energy of preferential interactions for SN−CCK4 as a reflection of the deviation of the experimental monomer−dimer midpoint from that of the stochastic distribution. Within the range of distributions shown by the shaded regions of Figure 3, the comparison demonstrates that the free energy of any preferential self-interactions between SN−CCK4 proteins in micelles is minimal and certainly less than one kcal mol^{-1}. By comparison with the calculated random distributions, it can be seen that the midpoints overlay almost exactly if the average aggregation number equals 50. In this case, we can conclude that the dimerization propensity of SN−CCK4s arises solely from the random distribution in micelles, e.g., association of SN−CCK4 is completely due to overloading of the micelles with protein and not from any preferential interactions between the helices. On the other hand, if the average aggregation number is 150, the midpoint of the experimental curve (1.5×10^{-2}) falls to the right of the stochastic curve (5×10^{-3}). This result might suggest the presence of slight repulsive interactions between the SN−CCK4 proteins. Physically, this is possible using this stochastic model, because there is a significant population of protein-free micelles at the midpoint of the random distribution. In this latter case, the mole fraction concentration difference between the midpoints of the random and experimental populations is approximately 3-fold. From this, we can estimate the free energy of repulsion as being on the order of the available thermal energy at 25 °C. Thus, even if with the uncertainty of knowing the size of the protein−micelle complex, we can conclude that the TMDs of CCK4, unlike those of glycophorin A, do not encode a thermodynamically meaningful preferential self-interaction in micelles.

Sequence Motifs Do Not Necessarily Encode Strong Self-Association. The lack of a preferential self-interaction between the helices was initially surprising to us in light of the evolutionary conservation of the transmembrane sequence, which includes a conserved GxxxG motif. Since it has been suggested that this pairwise glycine motif might be a sequence signature indicative of transmembrane α-helix dimerization (*12*), we hypothesized that the CCK4 TMDs would self-associate. The presence of a GxxxG motif is important for the most stable dimer formation of the glycophorin A TMD (*21*), and the GpA dimer has been invoked as a structural model for the way in which the glycines in a GxxxG motif might interact in a transmembrane dimer (*11*, *12*, *34*). Moreover, the leucine following the first glycine in the CCK4 GG4 motif was found to be the only stabilizing single point mutation in the context of the glycophorin A TMD (*21*), and we would therefore not predict that this small-to-large mutation would be inhibitory to protein interactions. However, a rationalization of the cumulative differences between the GpA and CCK4 transmembrane sequences is not straightforward, since it has been shown by double-mutant cycle analysis that the thermo-

dynamic interactions between the sites flanking and including the GpA GxxxG motif are cooperative and not additive (*22*). Glycophorin A TMD mutant sequences lacking a GxxxG motif can still dimerize, whereas other SN-GpA mutant sequences containing a GxxxG motif do not dimerize (*22*). As in glycophorin A, we hypothesize that the role of the GxxxG motif in the CCK4 TMD may be to allow a close approach of the helices in order for flanking sequences to interact. This might be important for disulfide bond formation in the full-length CCK4 protein, which contains a conserved cysteine in the flanking region C-terminal to the TMD. Alternatively, extramembranous regions of CCK4 may encode the protein−protein interactions for this receptor, and the TMD sequence need only permit a close approach of two or more proteins. The small amino acid glycine would facilitate this, and the remainder of the sequence need only be optimized to allow it.

In addition to a conserved GxxxG motif, TMSTAT analysis (*11*) of the CCK4 TMD reveals that it contains 25 pair wise motifs that are significantly overrepresented ($p < 0.05$) and six that are significantly underrepresented. Interestingly, only 31% of the overrepresented pairs involve evolutionarily conserved residues in the TMD, whereas this value is 50% for the underrepresented pairs. From the current data, we can conclude that the presence of a GxxxG and other pairwise motifs as arranged in the CCK4 TMD do not appear to be sufficient to encode strong helix−helix interactions. Additional experimentation is required to elucidate the functional importance of these motifs in transmembrane helices.

Functional Significance of the Absence of Preferential Self-Interactions between the CCK4 Transmembrane Domains. The lack a preferential self-association propensity for the CCK4 TMD may have several functional implications. First, we can postulate that the CCK4 TMD may play a relatively passive role in signaling, as long as the TMD sequence allows other regions of the full-length receptor to appropriately interact. Alternatively, the CCK4 TMD may encode an ability to interact with the TMDs of other receptor tyrosine kinases. The kinase domain of CCK4 has been shown to lack activity, suggesting that that the full-length CCK4 receptor may function like the erbB3 receptor tyrosine kinase, which also lacks kinase activity (*35*). It has been proposed that the evolutionary pressure for this lack of kinase activity might be that their cytoplasmic domains have extremely potent mitogenic power for recruiting downstream signaling molecules (*36*). Overactive signaling would be deleterious to cells if they also were able to transphosphorylate themselves in homodimeric complexes. Therefore, like erbB3, it is likely that signaling active complexes of CCK4 must be co-receptors containing another kinase-active partner. The TMD of CCK4 might be involved in mediating such hetero-interactions.

Alternatively, a second explanation for the lack of self-interactions interactions in the current experiments could be that that the functional properties of the TMD are not compatible with the micellar environment required for the sedimentation equilibrium analysis. In a cell, the biological lipid bilayer may impose additional structural and/or dynamic constraints upon the CCK4 TMD, which might stabilize conformations that would encode specific interactions in bilayers.

CONCLUSIONS

Regulated signaling of tyrosine kinase receptors is crucial for controlled cell growth. The observation that the CCK4 TMD self-associates much more weakly than the glycophorin A TMD is consistent with the current expectation that a tyrosine kinase receptor would not normally be constitutively found in a dimeric state. However, the complete lack of preferential self-interactions in micelles may suggest that the CCK4 TMD plays a supportive role in signaling. Future studies in lipid bilayers might possibly reveal specific interactions between the helices, and such a finding would emphasize the importance of the hydrophobic environment in modulating the structural and/or thermodynamic properties of this membrane protein. Alternatively, the conserved CCK4 TMD may mediate hetero-interactions. An understanding of the discrete stochastic distribution of membrane proteins in micelles and lipid bilayers is important for distinguishing between preferential and random interactions, both of which can influence the oligomeric population and perhaps the functional properties as well.

ACKNOWLEDGMENT

We thank Ann Marie Stanley for helpful discussions and critical reading of the manuscript.

REFERENCES

1. Prenzel, N., Fischer, O. M., Streit, S., Hart, S., and Ullrich, A. (2001) The epidermal growth factor receptor family as a central element for cellular signal transduction and diversification, *Endocr. Relat. Cancer 8*, 11−31.
2. Slamon, D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., and Ullrich, A. (1989) Studies of the HER-2/neu proto-oncogene in human breast cancer and ovarian cancer, *Science 244*.
3. Slamon, D., Clark, G., Wong, S., Levin, W., Ullrich, A., and McGuire, W. (1987) Human Breast Cancer: Correlation of Relapse and Survival with Amplification of the HER-2/*neu* Oncogene, *Science 235*, 177−182.
4. Easty, D. J., Mitchell, P. J., Patel, K., Florenes, V. A., Spritz, R. A., and Bennett, D. C. (1997) Loss of expression of receptor tyrosine kinase family genes PTK7 and SEK in metastatic melanoma, *Int. J. Cancer 71*, 1061−1065.
5. Radinsky, R., Risin, Fan, Dong, Bielenberg, Bucana, and Fidler. (1995) Level and function of epidermal growth factor receptor predict the metastatic potential of human colon carcinoma cells, *Clin. Cancer Res. 1*, 19−31.
6. Shaheen, R. M., Ahmad, S. A., Liu, W., Reinmuth, N., Jung, Y. D., Tseng, W. W., Drazan, K. E., Bucana, C. D., Hicklin, D. J., and Ellis, L. M. (2001) Inhibited growth of colon cancer carcinomatosis by antibodies to vascular endothelial and epidermal growth factor receptors, *Br. J. Cancer 85*, 584−589.
7. Mossie, K., Jallal, B., Alves, F., Sures, I., Plowman, G. D., and Ullrich, A. (1995) Colon carcinoma kinase-4 defines a new subclass of the receptor tyrosine kinase family, *Oncogene 11*, 2179−2184.
8. Chou, Y. H., and Hayman, M. J. (1991) Characterization of a member of the immunoglobulin gene superfamily that possibly represents an additional class of growth factor receptor, *Proc. Natl. Acad. Sci. U.S.A. 88*, 4897−4901.
9. Miller, M. A., and Steele, R. E. (2000) Lemon encodes an unusual receptor protein-tyrosine kinase expressed during gametogenesis in Hydra, *Dev. Biol. 224*, 286−298.
10. Pulido, D., Campuzano, S., Koda, T., Modolell, J., and Barbacid, M. (1992) Dtrk, a Drosophila gene related to the trk family of neurotrophin receptors, encodes a novel class of neural cell adhesion molecule, *EMBO J. 11*, 391−404.
11. Senes, A., Gerstein, M., and Engelman, D. M. (2000) Statistical analysis of amino acid patterns in transmembrane helices: the

GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions, *J. Mol. Biol. 296*, 921−936.
12. Russ, W. P., and Engelman, D. M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association, *J. Mol. Biol. 296*, 911−919.
13. Fleming, K. G., and Engelman, D. M. (2001) Specificity in transmembrane helix-helix interactions defines a hierarchy of stability for sequence variants, *Proc. Natl. Acad. Sci. U.S.A. 98*, 14340−14344.
14. Fleming, K. G. (2002) Standardizing the free energy change of transmembrane helix-helix interactions, *J. Mol. Biol. 323*, 563−571.
15. Fleming, K. G., Ren, C. C., Doura, A. K., Kobus, F. J., Eisley, M. E., and Stanley, A. M. (2004) Thermodynamics of glycophorin A transmembrane helix-helix association in C14 betaine micelles, *Biophys. Chem. 108*, 43−49.
16. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of Data from the Analytical Ultracentrifuge by Nonlinear Least-Squares Techniques, *Biophys. J. 36*, 575−588.
17. Laue, T. M., Shah, B., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided Interpretation of Analytical Sedimentation Data for Proteins, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., Horton, J. C., Eds.) pp 90−125, Royal Society of Chemistry, Cambridge.
18. Fleming, K. G., Ackerman, A. L., and Engelman, D. M. (1997) The Effect of Point Mutations on the Free Energy of Transmembrane α-Helix Dimerization, *J. Mol. Biol. 272*, 266−275.
19. Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B. J., Dempsey, C. E., and Engelman, D. M. (1992) Glycophorin A Dimerization is Driven by Specific Interactions between Transmembrane α-Helices, *J. Biol. Chem. 267*, 7683−7689.
20. Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J., and Engelman, D. M. (1992) Sequence Specificity in the Dimerization of Transmembrane α-Helices, *Biochemistry 31*, 12719−12725.
21. Doura, A. K., Kobus, F. J., Dubrovsky, L., Hibbard, E., and Fleming, K. G. (2004) Sequence context modulates the stability of a GxxxG mediated transmembrane helix-helix dimer, *J. Mol. Biol. 341*, 991−998.
22. Doura, A. K., and Fleming, K. G. (2004) Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer, *J. Mol. Biol. 343*, 1498−1497.
23. Sulistijo, E. S., Jaszewski, T. M., and MacKenzie, K. R. (2003) Sequence-specific dimerization of the transmembrane domain of the "BH3-only" protein BNIP3 in membranes and detergent, *J. Biol. Chem. 278*, 51950−51956.
24. Arkin, I. T., Adams, P. D., MacKenzie, K. R., Lemmon, M. A., Brunger, A. T., and Engelman, D. M. (1994) Structural organization of the pentameric transmembrane α-helices of phospholamban, a cardiac ion channel, *EMBO J. 13*, 4757−4764.
25. Lear, J. D., Gratkowski, H., and DeGrado, W. F. (2001) Membrane active peptides, *Biochem. Soc. Trans. 29*, 559−564.
26. Tanford, C., and Reynolds, J. A. (1976) Characterization of Membrane Proteins in Detergent Solutions, *Biochim. Biophys. Acta 457*, 133−170.
27. Hall, D., and Minton, A. P. (2003) Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges, *Biochim. Biophys. Acta 1649*, 127−139.
28. Lichtenberg, D., Robson, R. J., and Dennis, E. A. (1983) Solubilization of phospholipids by detergents. Structural and kinetic aspects, *Biochim. Biophys. Acta 737*, 285−304.
29. Neugebauer, J. (1990) *A Guide to the Properties and Uses of Detergents in Biology and Biochemistry*, Calbiochem Corporation, San Diego.
30. Møller, J. V., and le Maire, M. (1993) Detergent binding as a measure of hydrophobic surface area of integral membrane proteins, *J. Biol. Chem. 268*, 18659−18672.
31. le Maire, M., Champeil, P., and Møller, J. V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents, *Biochim. Biophys. Acta 1508*, 86−111.
32. le Maire, M., Kwee, K., Anderson, J. P., and Møller, J. V. (1983) *Eur. J. Biochem. 129*, 525−532.
33. Grefrath, S. P., and Reynolds, J. A. (1974) The molecular weight of the major glycoprotein from the human erythrocyte membrane, *Proc. Natl. Acad. Sci. U.S.A. 71*, 3913−3916.
34. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) The Calpha - - -H···O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions, *Proc. Natl. Acad. Sci. U.S.A. 98*, 9056−9061.
35. Mossie, K., Jallal, B., Alves, F., Sures, I., Plowman, G. D., and Ullrich, A. (1995) Colon carcinoma kinase-4 defines a new subclass of the receptor tyrosiine kinase family, *Oncogene 16*, 2179−2185.
36. Waterman, H., Alroy, I., Strano, S., Seger, R., and Yarden, Y. (1999) The C-terminus of the kinase-defective neuregulin receptor ErbB-3 confers mitogenic superiority and dictates endocytic routing, *EMBO J. 18*, 3348−3358.

BI048076L

doi:10.1016/j.jmb.2005.01.059 J. Mol. Biol. (2005) 347, 759–772

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

The Transmembrane Domains of ErbB Receptors do not Dimerize Strongly in Micelles

Ann Marie Stanley and Karen G. Fleming*

T.C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218 USA

*Corresponding author

The epidermal growth factor receptors (erbB) constitute an important class of single pass transmembrane receptors involved in the transduction of signals important for cell proliferation and differentiation. Receptor association is a key event in the signal transduction process, but the molecular basis of this interaction is not fully understood. Previous biochemical and genetic studies have suggested that the single transmembrane helices of these receptor proteins might play a role in stabilizing the receptor complexes. To determine if the erbB transmembrane domains could provide a driving force to stabilize the receptor dimers, we carried out a thermodynamic study of these domains expressed as C-terminal fusion proteins with staphylococcal nuclease. Similar fusion constructs have been used successfully to investigate the oligomerization and association thermodynamics of a number of transmembrane sequences, including that of glycophorin A. Using SDS-PAGE analysis and sedimentation equilibrium analytical ultracentrifugation, we do not find strong, specific homo or hetero-interactions between the transmembrane domains of the erbB receptors in micellar solutions. Our results indicate that any preferential interactions between these domains in micellar solutions are extremely modest, of the order of 1 kcal mol^{-1} or less. We applied a thermodynamic formalism to assess the effect of weakly interacting TM segments on the behavior of a covalently attached soluble domain. In the case of the ligand-bound EGFR ectodomain, we find that restriction of the ectodomain to the micellar phase by a hydrophobic TM, even in the absence of strong specific interactions, is largely sufficient to account for the previously reported increase in dimerization affinity.

Keywords: transmembrane dimer; membrane protein interactions; thermodynamics; erbB receptor; EGF receptor

Introduction

Transmission of molecular signals across biological membranes is essential for normal cellular function. The epidermal growth factor receptors represent an essential class of receptor tyrosine kinases (RTKs) involved in the transduction of signals important for cell proliferation and differentiation. In humans, this family includes four receptor proteins: the epidermal growth factor receptor (EGFR, erbB1, HER1), erbB2 (HER2), erbB3 (HER3), and erbB4 (HER4). The generally accepted paradigm for signaling by these receptors is a ligand-induced rearrangement of the extracellular domains followed by receptor association. The subsequent trans-phosphorylation of the intracellular kinase domains results in downstream signaling.[1,2] The nature of the signal is determined by the specific identity of the complex; these receptors are able interact with a number of hormone ligands to form a variety of homo and hetero-oligomeric complexes.[3] This diversity of interactions yields a vast signaling network, and

Abbreviations used: CAT, chloramphenicol acetyltransferase; C_8E_5, pentaethylene glycol monooctyl ether; EGFR, epidermal growth factor receptor; HER, human epidermal growth factor receptor; GpA, glycophorin A; RTK, receptor tyrosine kinase; SB-C14, 3-(*N*,*N*-dimethylmyristyl-ammonio)propanesulfonate; SDS, sodium dodecyl sulfate; SN, staphylococcal nuclease; SNGpA, staphylococcal nuclease/glycophorin A fusion protein; SNerbB, staphylococcal nuclease/erbB fusion protein; SRV, square-root of the variance; Strimer, separate trimer; TM, transmembrane; WT, wild-type.

E-mail address of the corresponding author: karen.fleming@jhu.edu

unregulated signaling within the network is implicated in a number of human cancers.[4]

Discovery of the molecular basis behind the protein association reaction in the erbB receptor complexes is critical to understanding the signal transduction process. Recent structures of the ligand-binding domains of the erbB receptors have begun to provide insight into the mechanisms of signaling. With the exception of erbB2, it is now clear that the extracellular domains (ECD) of these receptors exist in an auto-inhibited conformation in the absence of ligand. Instead of bridging two receptors, ligand-binding induces a conformational change in the ECD that reveals a cryptic protein–protein interface.[5] In contrast, the ECD of ErbB2, for which there is no known ligand, exists in an already extended conformation reminiscent of the structure induced in the ECD of the other erbBs by ligand-binding. Thus, the recent structures revealed that an interface between any two erbB proteins must be entirely receptor-mediated.[5]

While the crystal structures of the extracellular portions of the erbB proteins revealed a receptor–receptor interface, the extracellular ligand-binding domains alone are not capable of recapitulating all of the erbB interactions that are known to exist.[6,7] In some cases, the ligand-binding domain is not even necessary to form an active complex. The v-*erbB* oncogene product is a truncated homolog of erbB1 that lacks most of its ligand-binding domain, but is capable of dimerizing and transforming cells in a ligand-independent manner.[8,9] Additionally, a version of erbB2 with a truncated N-terminal domain also leads to more efficient transformation than its full-length counterpart, even though the ECD is presumably not normally auto-inhibited.[10,11] Interestingly, the extracellular domain of erbB2-(HER2) is proteolytically cleaved in tumor cells to yield a soluble fragment, detectable in the serum of breast cancer patients, and an active membrane fragment. The domain structure of the membrane fragment, with just a transmembrane segment and a tyrosine kinase domain, is reminiscent of the v-*erbB* oncogene product.[12,13] Together, these results suggest that other regions of the protein are important for stabilizing and/or regulating receptor–receptor interactions, and it raises questions about the relative contribution of the single transmembrane helix and the intracellular kinase domain to the receptor association.

A number of studies have suggested that specific interactions between the transmembrane helices are in fact important for stabilizing the receptor complexes. Attention first focused on the transmembrane domain as a potentially important site of interaction when the *neu* oncogene was identified as a point mutation of valine to glutamic acid within the transmembrane (TM) of a rat ortholog of erbB2.[14] Molecular modeling and solid-state NMR experiments suggest the glutamic acid might stabilize receptor dimers by forming inter-helical hydrogen bonds.[15,16] The transforming effect of the neu glutamic acid substitution was found to depend upon the surrounding sequence context,[17] as well as absolute position in the TM.[18] The surrounding sequence incorporates a five amino acid residue motif, Small-x-x-Large-G/A, consisting of a small residue (Gly, Ala, Ser, Thr, or Pro) in the zero position, a large aliphatic residue (Ala, Val, Leu, or Ile) in position 3, followed by Gly or Ala in position four.[15] This motif was identified in a large number of receptor tyrosine kinases, suggesting the motif could also be important for normal receptor function.[19] Recently, it was postulated that the two copies of the sequence motif in erbB1 have distinct functional roles, one being important for homo-interactions and the other important for hetero-interactions with erbB2.[20]

Additional evidence for a role for the transmembrane segment in receptor association resulted from a study of the ligand-induced dimerization of the extracellular domain of EGFR (erbB1). Extending this fragment to include the membrane-spanning segment enhanced the level of observed dimerization upon addition of the EGF ligand.[21] In other studies *in vivo*, peptides corresponding to the TM regions of the protein are capable of specifically inhibiting their cognate receptor.[22,23] Interfering with the transmembrane domain can also affect hetero-dimerization. Replacement of the transmembrane domain of erbB3 with the TM sequence from the fibroblast growth factor receptor or with a minimal lipid-anchor impairs hetero-dimerization with erbB2 relative to the wild-type erbB3, suggesting the TM segment functions as something more than a membrane tether.[6]

More recently, it was found that all four of the erbB TM segments gave strong positive responses in TOXCAT, a genetic assay for helix–helix association.[24] The erbB TM domains gave TOXCAT signals that were about half the strength of wild-type glycophorin A (GpA), a well-characterized TM model system known to have strong helix–helix interactions.[23,24] All four erbB TOXCAT responses were significantly stronger than that of the disruptive GpA mutant G83I.[24] This response level in TOXCAT can be interpreted as an indication of a high propensity for the transmembrane sequence to self-associate in bacterial membranes. Moreover, mutational analysis showed that changing residues in the Sternberg & Gullick motif[15,19] and the closely related GG_4 motifs[25] affected the response levels observed in TOXCAT.[24] Surprisingly, a valine to glutamic acid mutation in erbB2, analogous to that found in rat *neu*, which is generally believed to stabilize receptor dimers, resulted in a reduction in the TOXCAT signal.[24]

Whereas TOXCAT appeared to be reporting on a sequence-specific interaction, several studies indicate active receptors are actually quite promiscuous in the transmembrane sequences they can tolerate. EGFR can accommodate a number of proline substitutions in the TM and remain responsive to EGF.[26] Several erbB1 constructs with extended or truncated transmembrane domains will still bind ligand and retain their kinase activity.[26,27]

Additionally, the transmembrane domain of the rat neu receptor can be reduced to a simplified sequence of polyvaline with two appropriately placed glutamic acid residues without a loss of receptor dimerization or activation.[28,29] Thus, whether the transmembrane domains of the erbB receptors provide a strong and specific driving force for receptor interaction remains an open question.

In this study, we address the intrinsic propensity of the erbB transmembrane sequences to self-associate using a well-characterized fusion protein expression system coupled with sedimentation equilibrium analytical ultracentrifugation. We employed a construct in which the erbB transmembrane domains were fused *via* a 14 amino acid residue linker region to the C terminus of staphylococcal nuclease (SN). Similar SN-TM fusion constructs have been used to successfully probe the sequence-dependent oligomerization of several transmembrane domains, including glycophorin A, bNIP3, phospholamban, and several designed sequences.[30–33] Compared to these helix–helix oligomers, we find that the transmembrane domains of the erbB receptors have only a minimal propensity for interaction in micellar solutions, despite giving strong signals in TOXCAT.

Results

Table 1 shows the amino acid sequences for the TM segments of the four wild-type human erbB receptors that were used in this study. SNerbB2 was originally cloned without C-terminal charges, but subsequent inclusion of these charges in the SNerbB2KRR clone did not significantly impact the results. A construct in which the rat neu valine to glutamic acid mutation was mapped onto the human erbB2 sequence and another construct with a valine to isoleucine mutation that has been associated with a lower risk of cancer[34,35] were also cloned and expressed.

Strong dimerization cannot be detected in SDS micelles

The association of the SNerbB fusion constructs was first analyzed by SDS-PAGE. This method can detect stable oligomeric complexes of membrane proteins, and it has been used previously to map the interfaces of several TM helix oligomers.[30–32] Since SDS is a detergent micelle environment, membrane proteins may retain their structure, and some non-covalent interactions can be preserved. Indeed, the isolated transmembrane domain from the rat homolog of erbB2, protoneu, has been shown to retain a helical secondary structure in SDS micelles.[36]

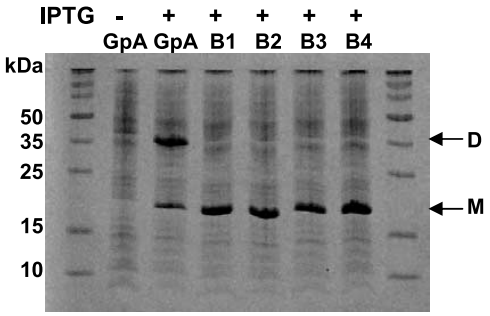


Figure 1. SDS-PAGE analysis of whole cell lysates for proteins expressing SN fusion proteins using a 20% acrylamide Phast gel. Arrows indicate the position to which monomers and dimers migrate. Lane 1, markers; lane 2, −IPTG SNGpA; lane 3, +IPTG SNGpA; lane 4, +IPTG SNerbB1; lane 5, +IPTG SNerbB2, lane 6, +IPTG SNerbB3; lane 7, +IPTG SNerbB4, lane 8, markers.

Figure 1 shows the SDS-PAGE of whole cell lysates, which revealed that the SNerbB fusion proteins migrated as monomers. For comparison, lysates from cells expressing SNGpA, which migrates predominantly as a dimer on SDS-PAGE, are also shown. Using purified protein, much higher concentrations (>100 μM) could be loaded onto the gel, but the protein still migrated as monomer (data not shown). Both SNerbB2 mutants, SNerbB2neu and SNerbB2VI, expressed as monomers, but the proteins were ill-behaved and neither could be stably purified in sufficient quantities for thermodynamic studies.

SNerbB proteins are predominantly monomeric in C_8E_5

In the case of the glycophorin A transmembrane domain, the free energy of the helix–helix interactions is known to be weaker in SDS relative to

Table 1. A list of the transmembrane sequences considered in this study, given by single-letter amino acid codes

Clone name	TM AA sequence
SNerbB1	SIATGMVGALLLLLVVALGIGLFMRRR
SNerbB2	LTSIVSAVVGILLVVVLGVVFGILI
SNerbB2KRR	LTSIVSAVVGILLVVVLGVVFGILIKRR
SNerbB2neu	LTSIVSAV**E**GILLVVVLGVVFGILIKRR
SNerbB2VI	LTSI**I**VSAV**V**GILLVVVLGVVFGILIKRR
SNerbB3	LTM**A**LTVIAGLVVIFMMLGGTFLYWRGRR
SNerbB4	LIAAGVIGGLFILVIVGLTFAVYVRRK

These sequences were fused to the C-terminal end of the staphylococcal nuclease *via* a linker region. Mutations are in bold and underlined.

other detergent environments.[37,38] Previously, sedimentation equilibrium analytical ultracentrifugation has been used to detect significant dimer populations for mutants of GpA that were monomeric by the SDS-PAGE assay.[39] In contrast to SDS-PAGE, analytical ultracentrifugation also allows for equilibrium measurements of the species in solution over a wide concentration range. This approach has been used to determine the thermodynamics of association of wild-type glycophorin A and more than 50 sequence variants.[39,40] Therefore, to more rigorously investigate the association of the erbB transmembrane segments, we carried out sedimentation equilibrium experiments in the neutrally buoyant detergent C_8E_5.

The SNGpA construct is greater than 90% dimeric in 23 mM C_8E_5 at the concentrations used in the centrifuge, and even the very disruptive GpA mutant G83I is greater than 20% dimeric under these conditions.[39] Therefore, we expected that we would be able to detect moderately stable interactions of the erbB TMs at this detergent concentration. Representative data sets for each SNerbB protein are shown in Figure 2. Using a non-linear, least-squares global analysis procedure, the data for each of SNerbB1, SNerbB2, and SNerbB3 were well described by a model for a single ideal species with the molecular mass of a monomer. Only SNerbB4 could not be described by a simple monomer model at 23 mM C_8E_5, as evidenced by the non-random residuals in Figure 2. However, if the concentration was raised to 50 mM C_8E_5, SNerbB4 was monomeric, suggesting the interactions in C_8E_5 are reversible. In contrast, most GpA mutants have a detectable population of dimer under these experimental conditions. These initial experiments suggest interactions between the erbB TMs are very weak, with SNerbB4 showing a slightly greater propensity for association relative to the other three.

At limiting protein to detergent ratios, SNerbB proteins self-associate

The apparent dissociation constant for dimerization of the GpA TM decreases as the detergent concentration is decreased, as a result of the reduced volume of the hydrophobic phase available to the protein.[41] Therefore, to populate the SNerbB4 oligomer(s), as well as possibly detect even weaker interactions for the other erbB TM domains, we decreased the detergent concentration in the sedimentation equilibrium experiments to 11 mM, a condition under which the disruptive glycophorin A mutant G83I would be greater than 50% dimeric.

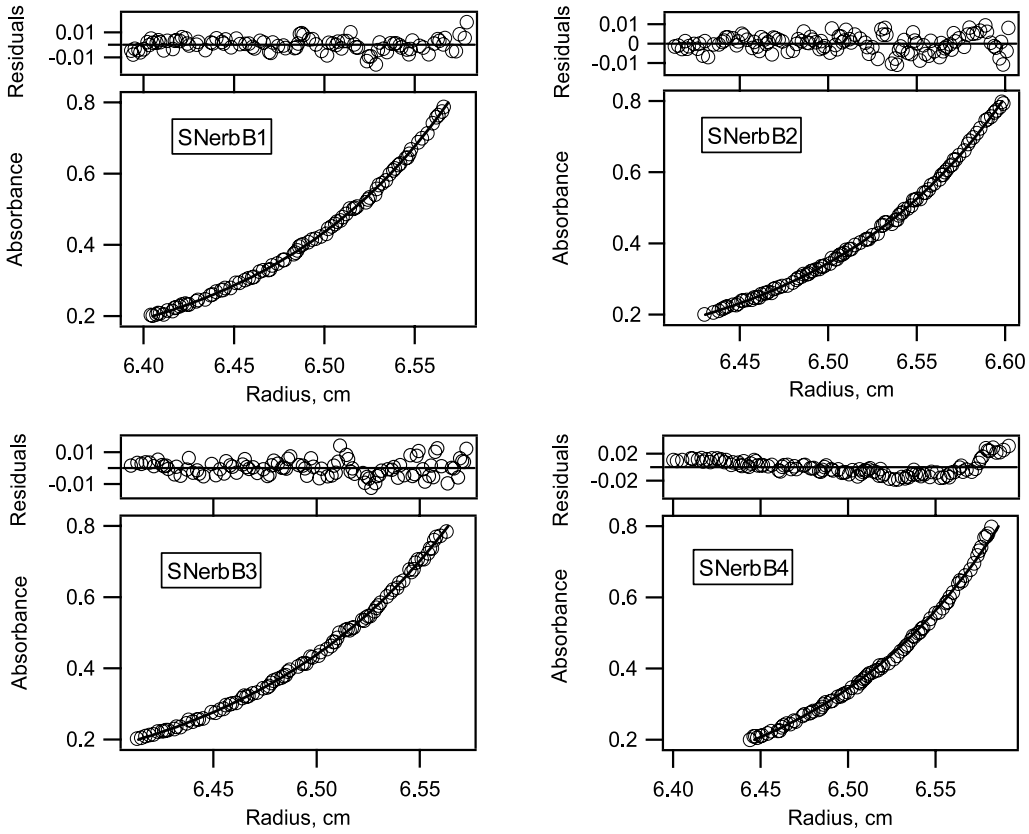


Figure 2. SNerbB proteins are monomeric at most detergent concentrations. Representative data sets from equilibrium AUC experiments are shown for the four erbB fusion constructs at 23 mM C_8E_5. The protein distribution was monitored by measuring absorbance at 280 nm. Raw data are shown as open circles; the continuous lines represent global fits for single ideal species with the molecular mass of a monomer. The upper panel for each data set shows the residuals for the fit. The residuals for the single ideal fit are small and random for SNerbB1, SNerbB2, and SNerbB3, indicating those data are well described by a monomer model. The residuals for SNerbB4, however, are larger and non-random.

A stochastic distribution model predicts some non-specific interactions at these high protein:detergent ratios used in the 11 mM C_8E_5 experiments,[42] but we sought to determine if we could drive the system to a preferential oligomer.

For all SNerbB constructs, the data at 11 mM were no longer described by a single monomeric species model and showed evidence for protein self-association. However, from these data alone, it was difficult to fit a specific association model (see discussion below), so we used the value for the reduced molecular mass, sigma (σ), as a model-independent parameter for expressing the extent of association, where σ is defined by Yphantis as:[43]

$$\sigma = \frac{M(1 - \bar{\upsilon}\rho)\omega^2}{RT} \quad (1)$$

where M is molecular mass, $\bar{\upsilon}$ is the partial specific volume of the protein (ml g^{-1}), ρ is the solvent density (g ml^{-1}), ω is the rotor speed (rad s^{-1}), R is the universal gas constant, and T is the temperature (K). For a single ideal species fit, the value of σ represents an approximate measure of the weight-average molecular mass of the species in solution. The expected value of σ for a given protein can be calculated from the amino acid sequence and the buffer components. If a protein is monomeric, then the experimental value for σ from a single ideal species model, σ_{exp}, would be equal to the calculated value, σ_{calc}. Thus, a ratio of $\sigma_{exp}/\sigma_{calc}$ greater than 1 indicates the presence of higher-order species, and larger ratios can indicate a greater extent of self-association. For SNerbB1, SNerbB2, and SNerbB3, the σ ratio is within error of unity at 23 mM C_8E_5, whereas the SNerbB4 ratio is significantly larger than 1 under the same conditions (Table 2). Even so, the experimental σ for SNerbB4 in 23 mM C_8E_5 is only 16% larger than the value calculated for monomer, indicating that monomer is still a significantly populated species. In contrast to the 23 mM data, Table 2 shows that the σ ratios for all four proteins are significantly greater than 1 when experiments are carried out in 11 mM C_8E_5. This result reflects the presence of higher-order species in all four proteins. The fact that the data at 11 mM C_8E_5 were no longer well described by a monomeric species alone is also reflected in the increase in the values of the square-root of the variance (SRV) for the single species fits as compared to the SRVs at 23 mM C_8E_5 (Table 2).

The absence of strong interactions was surprising, because the erbB TM domains show positive signals in the genetic assay TOXCAT, a result that is generally interpreted as an indication of strong helix–helix interactions.[23,24] In the TOXCAT assay, the activity of the erbB TMs was approximately half the activity of wild-type glycophorin A. Moreover, the erbB signals were several-fold greater than the disruptive glycophorin A mutant G83I. In contrast, in our experiments, the SNerbB proteins were monomeric under conditions where even G83I has a significant population of dimer.

Although the absolute strengths of the SNerbB interactions were much weaker than expected from the TOXCAT results by comparison to GpA, a consideration of just the relative signals of the erbBs suggests that the association trends in bacterial membranes as measured by TOXCAT are mirrored in the extent of association observed in the centrifuge in 11 mM C_8E_5. This comparison is shown in Figure 3, where it can be seen that in each assay the TM sequences of erbB1, erbB2, and erbB3 exhibit about the same levels of association, whereas there is evidence for a somewhat greater extent of association for erbB4. Although SNerbB4 exhibits a somewhat greater propensity for interaction than the other three sequences, the interactions of all four proteins in micellar solutions are very weak, especially by comparison to the GpA sequences.[39]

Specific association models for SNerbBs are difficult to define

We fit the sedimentation equilibrium data with a number of self-association models to determine the nature of species present in 11 mM C_8E_5. We tried a number of monomer–*N*mer models and found that for all four proteins a monomer–trimer model was the best description of the data from these simple two-species models. Table 3 shows the results of these fits for one of the experiments with SNerbB4, and results were similar for the other three proteins. The fact that a monomer–trimer equilibrium was the best fit to the data from the simple self-association models was surprising, since the

Table 2. Analysis of SNerbB AUC data using single ideal fits

		23 mM C_8E_5			11 mM C_8E_5		
SNerbB	σ_{calc}	SRV (10^{-3})	σ_{exp}	$\sigma_{calc}/\sigma_{exp}$	SRV (10^{-3})	σ_{exp}	$\sigma_{calc}/\sigma_{exp}$
1	0.9575	5.55	1.0036	1.0481	7.95	1.3657	1.4263
2	0.9287	6.71	0.9844	1.0600	8.18	1.2943	1.3937
3	0.9841	5.79	1.0581	1.0751	7.39	1.3593	1.3813
4	0.9533	7.79	1.1088	1.16311	9.06	1.6260	1.7057

At detergent concentrations of 23 mM or greater, the SNerbB proteins are monomeric in solution. When the detergent concentration is reduced to 11 mM, evidence of higher-order species can be seen. The values of σ for the four SNerbB fusion constructs were calculated using SEDNTERP for 20,000 rpm. σ_{exp} is the value returned by the NONLIN[47] fitting procedure for a single ideal species at 20,000 rpm. SRV is the square-root of the variance for these fits. The experimental numbers in the Table represent the average of at least two independent measurements.

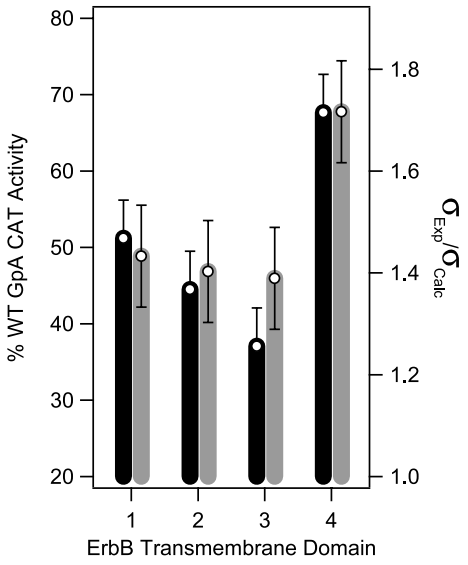
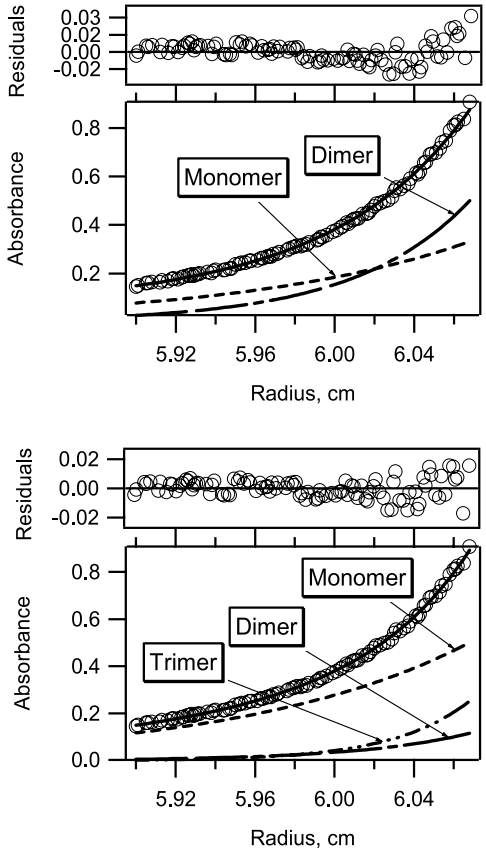


Figure 3. A comparison of the relative CAT activities of the erbB transmembrane sequences to the σ ratios from 11 mM C_8E_5 determined by sedimentation equilibrium using a single ideal species fit. It can be seen that in both assays, the transmembrane segment of erbB4 appears to be the most associated of the four transmembrane domains.

biology of the erbB receptors suggests that a monomer–dimer equilibrium would likely be the most appropriate model to describe the data.[44,45] Nevertheless, Figure 4 shows a comparison of the species plots for these fits, where it can easily be seen from the non-random residuals that a monomer–dimer equilibrium is not as good a description of the data as a monomer–trimer model.

Therefore, to independently identify the oligomers present in solution and to better discriminate between models for the sedimentation equilibrium data, we carried out glutaraldehyde cross-linking experiments. Cross-linking techniques can capture the oligomers that are preferentially populated in solution. The fusion proteins were cross-linked in 11 mM C_8E_5 at the highest protein concentrations used in the centrifugation experiments. Figure 5 shows that bands for both dimers and trimers could be seen, as well as traces of higher-order aggregates. At extended times, the SNerbB proteins all cross-link as large aggregates. The cross-linking results did confirm the presence of a dimer species, but they also revealed that it is likely at least three species are present in solution, which may be why a unique fit using the AUC data alone was difficult to define.

The cross-linking results and the biology of the



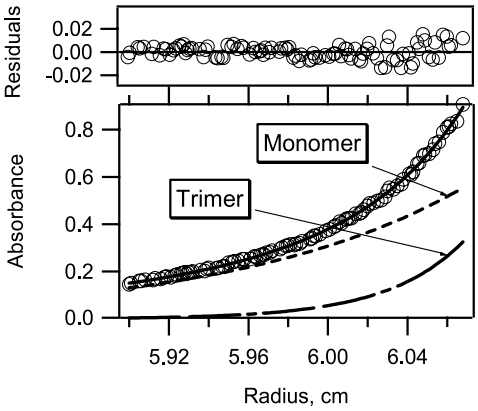


Table 3: Various association models fit to data for SNerbB4 in 11mmC_8E_5

Model	SRV
Single Ideal	8.56e-3
Monomer-Dimer	7.31e-3
Monomer-Trimer	6.49e-3
Monomer-Tetramer	7.92e-3
Monomer-Dimer-Trimer	6.42e-3
Monomer-DImer-Tetramer	6.45e-3

Figure 4. Species plots for various association models fit to data for SNerbB4 in 11 mM C_8E_5. Although only one data set is shown, a global fitting procedure using nine data sets, three concentrations at three rotor speeds was employed. It was expected that the equilibrium for SNerbB4 would be described by a monomer–dimer equilibrium, but higher order models appear to fit equally as well if not better, thus making it difficult to distinguish which is the appropriate description of the system.

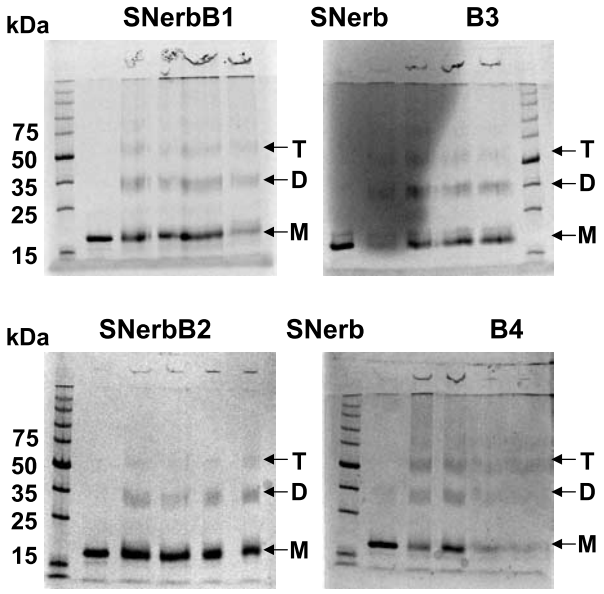


Figure 5. Higher-order SNerbB oligomers can be visualized by SDS-PAGE of glutaraldehyde cross-linked proteins. The proteins were cross-linked in 11 mM C_8E_5 at 43 μM with 0.1% glutaraldehyde and run on a 12.5% acrylamide Phast gel. Lane 1, molecular mass markers; lanes 2–6: 0, five, ten, 30, 90, and 120 seconds, respectively. Bands can be seen for both dimers and trimers; traces of higher-order aggregates can also be seen. At long times, all the proteins cross-link as large aggregates that crash out of solution.

erbBs justify the use of three species models that include dimers for the analysis of the sedimentation equilibrium data. Therefore, we fit the data with a monomer–dimer–trimer model, and the statistics of the fit and the shapes of the residuals indicated the model was a reasonable description of the data (Figure 4). Separation of the trimer terms, which can be interpreted as a non-equilibrating population for that species, led to an improvement in the fit. The inclusion of the additional parameters was supported by the Akaike information criterion[46] statistical test and analysis of the residuals by a runs test.[47]

Stochastic association largely accounts for the observed interactions

The observed species distributions for the SNerbB proteins in 11 mM C_8E_5 were compared to what is predicted by a purely stochastic distribution under these conditions. At these high protein to detergent ratios where the number of micelles approaches the number of protein molecules, the probability that more than one protein molecule can occupy a given micelle in the absence of a preferential interaction is non-negligible, and the distribution of proteins in micelles is given by a Poisson distribution as described in Materials and Methods.[42] Figure 6 shows a comparison of the experimentally observed species plots for SNerbB4 to those predicted by a stochastic model. The comparison shows that the fraction of monomer for SNerbB4 decreases more quickly than one would expect from purely stochastic interactions. It also demonstrates that the fraction of dimers expected from a random distribution is non-negligible at the concentrations used in the sedimentation equilibrium experiments, highlighted by the thickened portion of the line. The presence of yet another distinct species, namely the stochastic dimer, may also have contributed to the difficulty in defining a model for the AUC data at 11 mM C_8E_5. Attempts were made to fit the data with models that included a parameter to approximate the stochastic dimer, but this did not improve the fits. SNerbB4 exhibits the greatest propensity of the erbBs to associate, so the curves for SNerbB1, SnerbB2, and SNerbB3 were shifted even closer to the random distribution curves. They do remain to the left of the simulated data, indicating that they are slightly more associated than one would expect from a purely random distribution. However, we estimate that any preferential interactions between the erbB TMs are extremely modest, certainly less than 1 kcal mol^{-1}, and we conclude that stochastic interactions account for a large portion of the SNerbB association seen in 11 mM C_8E_5. It is worth noting though that at 23 mM C_8E_5, the stochastic model predicts a negligible amount of association but SNerbB4 is not monomeric under these conditions.

Heterodimers of the SNerbB proteins could not be detected

Since interactions in the transmembrane domains have been implicated in the heterodimerization of the erbB receptor tyrosine kinases,[6,20] we carried out experiments to detect hetero-interactions of the SNerbB fusions using SDS-PAGE. For each of the six pairwise interactions, equal amounts of both proteins were combined, but only monomer was visible on the gel when the samples were analyzed by SDS-PAGE (data not shown). We also attempted to detect modest interactions by sedimentation equilibrium between SNerbB2 and SNerbB3. We chose this pair because the erbB2/erbB3 heterodimer is the most potent signaling complex[48,49] and only weak hetero-interactions can be detected between the soluble ligand-binding domains,[7] so we hypothesized that if TM–TM interactions were important for stabilizing any of the hetero-complexes it would be the erbB2/erbB3 complex. Our results, however, showed that an equimolar amount of SNerbB2 and SNerbB3 at 23 mM C_8E_5 yielded a molecular mass consistent with protein monomers in solution (Figure 7).

Discussion

ErbB TM interactions are difficult to detect in micellar solutions

Transmembrane helix association as a driving

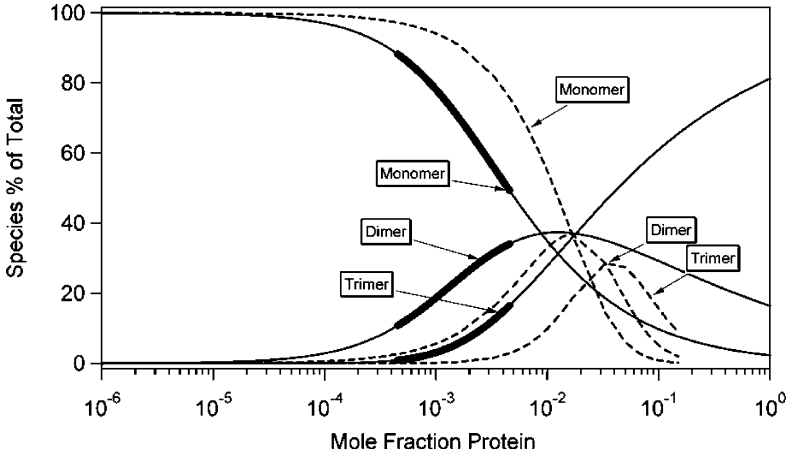


Figure 6. An overlay of a stochastic distribution and the distribution calculated from a monomer–dimer–separate trimer fit to one of the SNerbB4 data sets. The dotted lines represent the monomer, dimer, and trimer populations simulated using a random distribution of proteins into micelles for an aggregation number of 60 and a maximum occupancy of $k=5$. The continuous lines represent the distributions calculated using the global K_{12} constant and the K_{13} from the data set for the middle concentration at 24,500 rpm for 11 mM C_8E_5. The thickened black portion of the line represents the concentration region where the data were actually collected. It can be seen that the association of SNerbB4 is only very moderately shifted from what one would expect from a stochastic distribution.

force for protein association is not unprecedented, and several TM segments have been identified that are capable of driving oligomerization.[30–32] Much consideration has been given as to whether this is the role of the erbB transmembrane domains. Some studies have indicated that active EGF receptors can tolerate a variety of transmembrane sequences,[26–29] but others studies have suggested a specific role for the transmembrane domain in dimerization, with perhaps the most suggestive evidence for a role for TM helix interaction coming from the genetic assay TOXCAT.[15,21,24]

Contrary to what was expected from the TOXCAT,[24] we found that the transmembrane domains of the erbB receptors were not sufficient to drive the strong dimerization of an SN fusion partner in detergent micelles. This distinguishes the erbB TM sequences from the transmembrane domains of GpA and bNIP3, which are capable of driving robust dimerization of chimeric proteins in a number of environments, including detergent micelles and bacterial membranes.[30–32,37,39,50–52] It was also initially surprising that the erbB TMs did not interact, given that the sequences contain GxxxG motifs, which can be an important sequence motif in strong helix–helix interactions.[25,31] However, it has been subsequently shown that a GxxxG motif is not sufficient to drive TM helix dimerization.[39,42] Our results are also consistent with an earlier report of a similar SN/erbB1 TM fusion construct migrating as a monomer on SDS-PAGE.[53]

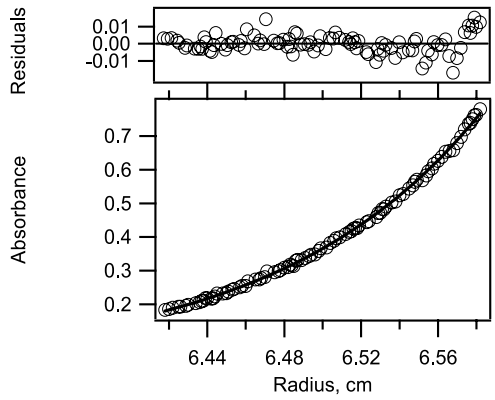


Figure 7. A representative sedimentation equilibrium data set for a mixture of equimolar amounts of SNerbB2 and SNerbB3 fit to a single ideal species. The raw data are represented as open circles and the continuous line represents the fit and the upper panel shows the residuals for the fit. The data are well described by a single ideal species with a molecular mass of 21,273 kDa, which is equal to the average of the two monomers, SNerbB2 (20,876 kDa) and SNerbB3 (21,669 kDa).

It is possible that the lack of interaction is a consequence of the SN fusion construct that was used, but we believe this is unlikely. Our linker region is 14 amino acid residues long and is comparable to what has been used in the past to successfully measure the energetics of glycophorin A, a number of GpA mutants, and other TM helices.[31,33,37,39,40] In all of these studies, the SN fusion did not appear to interfere, and the linker length was sufficient to allow the TMs to interact. Additionally, extending the linker between the SN domain and the TM sequence had no effect on the interaction behavior of the CCK-4 TM domain.[42] Furthermore, it has been reported that a peptide corresponding to the TM of the EGF receptor (erbB1) migrated as a monomer in SDS-PAGE,[54] and isolated erbB2 TM peptides also migrate predominantly as monomers.[55] These results with isolated TM peptides are consistent with the SDS-PAGE analysis of our SN-fusion constructs. Some reports do indicate, however, peptides corresponding to the TM of the proto-oncogenic neu receptor from rats, a homolog of erbB2, migrate as something larger than monomers on SDS-PAGE.[36,55,56]

While specific protein–protein interactions determined by the amino acid sequence are important for TM helix interactions, the environment can also be a critical factor in the investigation of membrane proteins. Successful solution studies can depend on the choice of an appropriate detergent. However, our inability to detect strong interactions does not

seem to be simply the effect of the selected experimental environment, as the protein did not interact in a number of detergents. The SNerbB proteins are principally monomeric in the non-ionic detergent C_8E_5, as well as in the ionic detergent SDS. Preliminary sedimentation equilibrium experiments in the zwitterionic detergent SB-C14 also failed to detect strong SNerbB interactions (data not shown).

Detergent micelles have proven to be powerful and useful tools in many membrane protein studies, but detergent micelles are sometimes criticized as being poor substitutes for the two-dimensional fluid of a lipid bilayer. In the case of the erbB proteins, it may be that the bilayer imposes structural constraints that favor dimerization-competent conformations of the erbB transmembrane domains but these constraints would be lacking in a micellar environment. The results from the TOXCAT assay, which is carried out in bacterial membranes, are certainly consistent with this idea.[24] However, the populations in solid-state NMR experiments in model bilayers seem to be in agreement with the level of association we observe. While the solid-state NMR experiments are not a direct measure of association and can be difficult to interpret, these studies suggest that the erbB1 TM is largely monomeric in bilayers, with the only signs of interactions being seen at very high concentrations of protein.[57–61] The erbB2 TM appears to exhibit a somewhat greater tendency to associate in the solid-state NMR experiments, but there is still a significant population of molecules that behave as one would expect for an unrestricted monomer.[61] These NMR experiments suggest that the erbB TM peptides are not constitutively associated in bilayers, and thus our inability to detect strong interactions between the erbB TM domains may not be the consequence of the micellar environment required for the sedimentation equilibrium studies. We cannot, however, eliminate the possibility that interactions we detect as extremely weak in solution are enhanced by the environment of the native bilayer and are biologically relevant.

Another possible explanation for the apparent discrepancy between TOXCAT and the other assays for helix–helix interactions is the difference in the domains flanking the transmembrane segment. In the full-length receptors, the transmembrane helix connects a ligand-binding ectodomain and an intracellular kinase domain. It may be that the TM segments of the erbB receptors cannot function as independent interaction domains isolated from the context of the full-length protein. This deficiency may somehow be compensated for by the nature of the soluble domains in the TOXCAT chimera. Of the studies considered here, the TOXCAT construct is the only construct that has a soluble domain with a dimerization interface, in the form of the ToxR DNA-binding domain. The TOXCAT construct is also the only construct with a C-terminal soluble domain. It may be that the erbB TMs require interactions in the soluble domains to pre-arrange the TM helices and that the TOXCAT fusion protein has an architecture to accomplish this. It could also be that the presence of a structured C-terminal domain is important for stabilizing interactions at the end of the transmembrane segment necessary for the receptor association. Previously, it has been suggested that the erbB transmembrane helix and the intracellular kinase domain are rotationally coupled, implying some degree of rigidity in the linker between the two.[28] It is also worth noting that this C-terminal juxtamembrane region of EGFR contains an important phosphorylation site for protein kinase C[62,63] and that the addition of peptides corresponding to this juxtamembrane region leads to increased receptor autophosphorylation.[64]

Consequences of including a TM lacking strong specific self-associations

In studies of SNGpA fusion proteins, it has been shown that it is appropriate to treat the micelles as a distinct thermodynamic phase and to consider the transmembrane segments confined to this phase.[41,65,66] Such a consideration leads to the expression:[41]

$$\Delta G_{\text{app}} = \Delta G_x^0 + RT \ln[\text{Micellar Det}]_{\text{w}} \qquad (2)$$

where ΔG_{app} is the observed apparent free energy in bulk aqueous concentrations, ΔG_x^0 is the standard state mole fraction free energy of association, R is the universal gas constant, T is the temperature (K), and $[\text{Micellar Det}]_{\text{w}}$ is the aqueous concentration of the detergent in micelles. The first term represents the association free energy of the intrinsic chemical components of the association reaction, and the second term represents a statistical energy of mixing, the magnitude of which varies with the micellar detergent concentration. In the case of GpA, we believe the contribution of the soluble SN domain to the standard state free energy ΔG_x^0 is negligible, and we interpret ΔG_x^0 as the contribution from interactions between the transmembrane domains.

However, the soluble domain of EGFR, when ligand-bound, does interact. Equation (2) predicts that extending these fragments to include a hydrophobic transmembrane segment should enhance the apparent dimerization affinity of the ECDs even in the absence of specific TM interactions. The TM domain would be restricted to the volume of the micellar phase, and thus the apparent free energy of association of the ECDs would now contain a contribution from the statistical energy term described above as well as the intrinsic free energy of dimerization of the soluble domains, ΔG_{int}. Applying the above formalism, we can substitute ΔG_{int} for ΔG_x^0 and rearrange equation (2) to write the apparent association equilibrium constant for an ECD-TM construct in micelles, K_{app}, as:

$$K_{\text{app}} = \frac{K_{\text{int}}}{[\text{Micellar Det}]_{\text{w}}} \qquad (3)$$

The equilibrium constants in equation (3) are in bulk molar concentration units. From equation (3), one would expect that the association of the receptor would depend on detergent concentration, with a tighter apparent association expected at lower detergent concentrations, and it has in fact been shown that the full-length EGFR, as well as a truncated fragment resembling v-erbB, interact more strongly at low detergent concentrations.[67,68]

A study by Tanner & Kyte, which compared the ligand-induced dimerization of the ECD of erbB1 to the dimerization of a fragment extended to include the TM, is also consistent with this formalism. This study found the longer fragment did indeed associate more readily, although complete dimerization was never observed.[21] The dimerization of the longer fragment was measured in 1% Triton X-100 (15 mM), and equation (3) predicts that restriction of the soluble fragment to the micellar phase would lead to an apparent 67-fold enhancement in the association. By comparing the dimerization of the ECD alone, which had been shown to be complete at 40 μM protein,[69] with the dimerization of the ECD-TM fragment in Triton micelles that they observed (2 nM), the authors estimated an observed enhancement of approximately 10,000-fold.[21] However, a subsequent study reports complete ligand-induced dimerization of the extracellular domain of erbB1 at concentrations as low as 4 μM protein;[7] therefore, the enhancement of ligand-induced dimerization of the ECD by the addition of the TM may be less than originally suggested and be more of the order of 1000-fold. Equation (3) suggests a large portion of this is simply the result of restricting the longer fragment to be stably inserted in the micelles. If approximately 65-fold arises from limiting the protein to the micellar phase, only a 15–30-fold enhancement arises from other features of the ECD-TM construct. This additional enhancement is modest, and it could arise from stabilization in the longer construct of structures in the juxtamembrane region important for dimerization or from restriction of the relative orientation of two ECDs on the surface of a micelle. The role of the transmembrane domain in the differences observed in the Tanner & Kyte study is still unclear, but our results suggest that it may not be necessary to invoke strong specific interactions between the TM domains to account for the observed enhancement.

A sterically permissible hydrophobic transmembrane segment may be sufficient

Some studies have indicated active receptors tolerate a good deal of sequence promiscuity in the transmembrane segment,[26,29] but others have implicated sequence specificity in the TM is necessary to maintain proper receptor interactions in the EGFR family.[17,18,23] If there are not strong interactions between the transmembrane domains, then why do some TM sequence variants affect receptor function? This dual behavior might be reconciled in a case where proper orientation of the soluble domains puts the transmembrane domains in close proximity without the transmembrane segments necessarily providing significant energy to stabilize the dimer. In other words, it may only be required that the transmembrane region dimers be sterically permissible. Examination of the crystal structures of the extended ligand-bound conformation of the erbB extracellular domains and structure-based models of the receptor dimers reveal that the C termini of domain IV point toward one another in way that could juxtapose the transmembrane domains of two receptors.[70–73] The need for a close-approach of the transmembrane domains might explain the occurrence of motifs that place small residues on the same face of the helix, such as the Stenberg & Gullick and GG_4 motifs.

The receptors could be destabilized by mutations in the TM that introduce large steric clashes in regions of the transmembrane segments positioned close in space by the orientation of the soluble domains. For mutations to have this effect, there would have to be some degree of rigidity in the structure so that the soluble domains and the mutated transmembrane sequence could not simultaneously adopt their optimal conformations. A previous study that explored the rotational coupling between the transmembrane segment and the kinase domain of the neu receptor has in fact suggested that the linker between the two domains is rigid.[28] Conversely, sequence variants that could be juxtaposed when the soluble domains of the protein adopt their preferred conformation in the dimer would be tolerated.

Such a situation would be consistent with what is observed for the neu mutation. The wild-type valine residues may be in close proximity in the transmembrane region, but not interacting to any significant extent. Placing a glutamic acid residue in an already optimal position introduces a hydrogen-bonding interaction, in an otherwise non-interacting domain, which stabilizes the dimer. The subsequent introduction of an unfavorable interaction, such as one might expect if glycine residues in close proximity are mutated to larger residues, would then hinder the receptor association.[17] When placed in a non-optimal position elsewhere in the transmembrane domain, no effect is observed.[18] It may be the case that the sterics of the soluble domains prevent the glutamic acid residues in the transmembrane domains from interacting. Alternatively, it may be an effect of the sequence context of the TM itself.[74] Since a number of studies have shown that the introduction of a single appropriately placed polar residue in a TM can lead to association of an otherwise non-interacting TM domain,[33,75,76] we do not believe that it is a prerequisite for the wild-type erbB transmembrane domains to have significant interactions in order for the neu mutation to effect cellular transformation by stabilization of receptor dimers.

Conclusions

Contrary to what was expected, given the behavior of the erbB transmembrane segments in TOXCAT, we do not detect strong specific interactions between the TM segments of the erbB proteins in detergent micelles using an SN-TM fusion system. This is in contrast to the transmembrane domains of GpA and bNIP3, which give strong TOXCAT signals and are capable of driving stable protein oligomerization in a number of environments. Our study suggests that one interpretation of the results is the transmembrane segments of the erbB proteins isolated from the full-length protein are not independent association domains and cannot drive oligomerization, at least in a micellar environment. It may be that interactions we detect in solution as extremely modest are significantly enhanced in a bilayer environment. However, given that the biology of the receptor dictates that protein exists in a monomer–dimer equilibrium, it may not be surprising that TM interactions are distinctly weaker than for those of glycophorin A. Even a non-interacting hydrophobic transmembrane domain, however, can increase the apparent affinity of the covalently attached soluble domain by restricting the protein to the hydrophobic phase of the solution. The role of the transmembrane domain in the context of the full-length protein remains unclear, but it may be sufficient for the transmembrane sequences of the erbB proteins to be sterically permissible. In this scenario, the transmembrane domains are passive, in the sense that they do not contribute significant energy; however, not just any sequence will suffice because the packing of the TM segments must not interfere with the association of the other domains. The contribution of the individual domains of the receptors to the full-length complexes is unclear, but it seems likely the association of the erbB receptors will be the sum of many interactions, including numerous weak contacts, distributed throughout all the domains.

Materials and Methods

Cloning and expression

Quikchange mutagenesis was used to introduce a SmaI site into the pet11b-SNGpA plasmid N-terminal to the GpA TM domain. The GpA TM sequence was excised from the SN vector containing an open reading frame for a staphylococcal nuclease C-terminal GpA fusion by restriction digest with either SmaI/BamHI or XmaI/BamHI, leaving behind the reading frame for staphylococcal nuclease and a C-terminal linker. The erbB2 transmembrane sequence was generated by a PCR of a pccB2 plasmid template, a generous gift from Dr Mark Lemmon and Dr Jeanine Mendrola, using primers containing SmaI and BglII restriction sites. BglII creates an overhang compatible with BamHI and the erbB2 TM sequence contains an internal BamHI site. The PCR product was digested and ligated into the SN vector. The other erbB sequences were generated by ordering synthetic 5′ phosphorylated oligonucleotides with the appropriate overhangs from Invitrogen and annealing the oligonucleotides by heating to 95 °C with subsequent cooling. The annealed oligonucleotides were then ligated into the SN vector. SnerbB2 variants were made using the Stratagene Quikchange protocol, and the sequences for all constructs were verified by DNA sequencing. The SNerbB plasmids were transformed into HMS174(DE3) cells, and expression of the protein was induced with IPTG. The SNerbB proteins were purified using Thesit® detergent according to the protocol published for SNGpA.[30,37]

SDS-PAGE assay

Cells were grown to an A_{600} of 0.8, induced with 1 mM IPTG for four hours, and harvested by centrifugation. Cell pellets were resuspended in SDS-PAGE loading buffer and analyzed using a 20% (w/v) acrylamide gel using a Pharmacia Phast gel system. For hetero-dimer interactions, equal amounts of protein at 25 μM in 20 mM NaPi 7.0, 100 mM NaCl, and 2% (w/v) Thesit® were mixed and incubated for 30 minutes at 37 °C. The samples were then analyzed on a 20% acrylamide Phast gel.

Sedimentation equilibrium analytical ultracentrifugation

Purified proteins were exchanged into the neutrally buoyant C_8E_5 detergent by ion-exchange chromatography as described.[37] Sedimentation equilibrium experiments were performed in Beckman XL-A analytical ultracentrifuge in six-sector cells according to the previously published protocol for SNGpA, with the exception that the protein was observed at 280 nm.[37] Data were collected at three initial protein concentrations (A_{280}=0.9, 0.6, 0.3) and three rotor speeds (20,000 rpm, 24,500 rpm, and 30,000 rpm). WINMATCH† was used to check that equilibrium was reached at each speed. For the data analysis, protein molecular masses and partial specific volumes were calculated using the program SEDNTERP.[77] Since none of the transmembrane sequences of interest contained tryptophan, the absorbance of the TM was considered negligible relative to the SN part of the fusion and an extinction coefficient of A_{280}=1 mg/ml was used. The radial distribution profiles for the nine data sets were globally fit by a non-linear, least-squares curve-fitting procedure in the Windows version of NONLIN.[47] The Akaike information criterion (AIC) was calculated using the equation:[46]

$$\mathrm{AIC} = -N \ln \sigma^2 - 2P \quad (4)$$

where N is the number of degrees of freedom, σ is the square-root of the variance, and P is the number of parameters in the model. The runs test score was calculated for each data set from the residuals of the global fit using the equations described earlier.[78]

Glutaraldehyde cross-linking

Glutaraldehyde (1%, w/v) in buffer was added to 40 μM protein in 11 mM C_8E_5, 20 mM NaPi 7.0, and 200 mM NaCl to a final concentration of 0.1% glutaraldehyde. Aliquots were taken at five, ten, 30, 90, and 120 seconds and quenched with SDS-PAGE loading buffer.

† www.bbri.org/rasmb

Samples were boiled for five minutes and analyzed on 12.5% acrylamide Phast gels.

Stochastic distribution data simulation

When the number of micelles and the number of protein monomers far exceeds the occupancy of a micelle, the distribution of the protein into the micelles due to random chance can be described by a Poisson distribution:

$$\mathrm{Lim}(m, n>>k) : \mathrm{Prob}[X_i = k] = \frac{1}{k!}\left(\frac{m}{n}\right)^k \mathrm{e}^{-m/n} \quad (5)$$

where X_i is the random variable that counts the number of proteins in a micelle i, k is the number of proteins occupying a micelle, m is the number of protein molecules, and n is the number of micelles.[42] Given a stochastic distribution, the fraction of proteins that are in a particular oligomeric state, f_j where j is a particular value of the occupancy, can be calculated from the following partition function, which considers only occupied micelles:

$$f_j = \frac{j\,\mathrm{Prob}[X_i = j]}{\sum_{k>0}^{k=5} k\,\mathrm{Prob}[X_i = k]} \quad (6)$$

The series was truncated at an occupancy of $k=5$ because oligomers larger than this did not contribute significantly to the distribution at micelle to protein ratios smaller than 1:1. Additionally, oligomers larger than pentamer would spin out at the speeds used in the sedimentation equilibrium experiments and are unobservable in the centrifuge. To calculate the species distribution as a function of total micellar detergent instead of the concentration of micelles, the ratio m/n was divided by the aggregation number. An aggregation number of 60, which lies between the aggregation numbers reported for C_8E_4 and C_8E_6,[79] was used.

Acknowledgements

This work was supported by Career and Idea awards from the Department of Defense (DAMD17-02-1-0427). A.M.S. is a Howard Hughes Institute Predoctoral Fellow. We gratefully thank Dr Mark Lemmon and Dr Jeannine Mendrola for the generous gift of the pccB2 plasmids and Maridel Lares for technical assistance. We thank Dr Dan Leahy for helpful comments on the manuscript.

References

1. Schlessinger, J. (2002). Ligand-induced, receptor-mediated dimerization and activation of EGF receptor. *Cell*, **110**, 669–672.
2. Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. *Cell*, **103**, 211–225.
3. Riese, D. J. 2nd & Stern, D. F. (1998). Specificity within the EGF family/ErbB receptor family signaling network. *Bioessays*, **20**, 41–48.
4. Mosesson, Y. & Yarden, Y. (2004). Oncogenic growth factor receptors: implications for signal transduction therapy. *Semin. Cancer Biol.* **14**, 262–270.
5. Burgess, A. W., Cho, H. S., Eigenbrot, C., Ferguson, K. M., Garrett, T. P., Leahy, D. J. *et al*. (2003). An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Mol. Cell*, **12**, 541–552.
6. Tzahar, E., Pinkas-Kramarski, R., Moyer, J. D., Klapper, L. N., Alroy, I., Levkowitz, G. *et al*. (1997). Bivalence of EGF-like ligands drives the ErbB signaling network. *EMBO J.* **16**, 4938–4950.
7. Ferguson, K. M., Darling, P. J., Mohan, M. J., Macatee, T. L. & Lemmon, M. A. (2000). Extracellular domains drive homo- but not hetero-dimerization of erbB receptors. *EMBO J.* **19**, 4632–4643.
8. Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P. *et al*. (1984). Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. *Nature*, **307**, 521–527.
9. Adelsman, M. A., Huntley, B. K. & Maihle, N. J. (1996). Ligand-independent dimerization of oncogenic v-erbB products involves covalent interactions. *J. Virol.* **70**, 2533–2544.
10. Di Fiore, P. P., Pierce, J. H., Kraus, M. H., Segatto, O., King, C. R. & Aaronson, S. A. (1987). erbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. *Science*, **237**, 178–182.
11. Segatto, O., King, C. R., Pierce, J. H., Di Fiore, P. P. & Aaronson, S. A. (1988). Different structural alterations upregulate *in vitro* tyrosine kinase activity and transforming potency of the erbB-2 gene. *Mol. Cell. Biol.* **8**, 5570–5574.
12. Molina, M. A., Codony-Servat, J., Albanell, J., Rojo, F., Arribas, J. & Baselga, J. (2001). Trastuzumab (herceptin), a humanized anti-Her2 receptor monoclonal antibody, inhibits basal and activated Her2 ectodomain cleavage in breast cancer cells. *Cancer Res.* **61**, 4744–4749.
13. Christianson, T. A., Doherty, J. K., Lin, Y. J., Ramsey, E. E., Holmes, R., Keenan, E. J. & Clinton, G. M. (1998). NH2-terminally truncated HER-2/neu protein: relationship with shedding of the extracellular domain and with prognostic factors in breast cancer. *Cancer Res.* **58**, 5123–5129.
14. Bargmann, C. I., Hung, M. C. & Weinberg, R. A. (1986). Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell*, **45**, 649–657.
15. Sternberg, M. J. & Gullick, W. J. (1989). Neu receptor dimerization. *Nature*, **339**, 587.
16. Smith, S. O., Smith, C. S. & Bormann, B. J. (1996). Strong hydrogen bonding interactions involving a buried glutamic acid in the transmembrane sequence of the neu/erbB-2 receptor. *Nature Struct. Biol.* **3**, 252–258.
17. Cao, H., Bangalore, L., Bormann, B. J. & Stern, D. F. (1992). A subdomain in the transmembrane domain is necessary for p185neu* activation. *EMBO J.* **11**, 923–932.
18. Bargmann, C. I. & Weinberg, R. A. (1988). Oncogenic activation of the neu-encoded receptor protein by point mutation and deletion. *EMBO J.* **7**, 2043–2052.
19. Sternberg, M. J. & Gullick, W. J. (1990). A sequence motif in the transmembrane region of growth factor receptors with tyrosine kinase activity mediates dimerization. *Protein Eng.* **3**, 245–248.
20. Gerber, D., Sal-Man, N. & Shai, Y. (2004). Two motifs within a transmembrane domain, one for homodimerization and the other for heterodimerization. *J. Biol. Chem.* **279**, 21177–21182.
21. Tanner, K. G. & Kyte, J. (1999). Dimerization of the extracellular domain of the receptor for epidermal

growth factor containing the membrane-spanning segment in response to treatment with epidermal growth factor. *J. Biol. Chem.* **274**, 35985–35990.
22. Lofts, F. J., Hurst, H. C., Sternberg, M. J. & Gullick, W. J. (1993). Specific short transmembrane sequences can inhibit transformation by the mutant neu growth factor receptor *in vitro* and *in vivo*. *Oncogene*, **8**, 2813–2820.
23. Bennasroune, A., Fickova, M., Gardin, A., Dirrig-Grosch, S., Aunis, D., Cremel, G. & Hubert, P. (2004). Transmembrane peptides as inhibitors of ErbB receptor signaling. *Mol. Biol. Cell*, **15**, 3464–3474.
24. Mendrola, J. M., Berger, M. B., King, M. C. & Lemmon, M. A. (2002). The single transmembrane domains of ErbB receptors self-associate in cell membranes. *J. Biol. Chem.* **277**, 4704–4712.
25. Russ, W. P. & Engelman, D. M. (2000). The GxxxG motif: a framework for transmembrane helix–helix association. *J. Mol. Biol.* **296**, 911–919.
26. Carpenter, C. D., Ingraham, H. A., Cochet, C., Walton, G. M., Lazar, C. S., Sowadski, J. M. *et al*. (1991). Structural analysis of the transmembrane domain of the epidermal growth factor receptor. *J. Biol. Chem.* **266**, 5750–5755.
27. Kashles, O., Szapary, D., Bellot, F., Ullrich, A., Schlessinger, J. & Schmidt, A. (1988). Ligand-induced stimulation of epidermal growth factor receptor mutants with altered transmembrane regions. *Proc. Natl Acad. Sci. USA*, **85**, 9567–9571.
28. Bell, C. A., Tynan, J. A., Hart, K. C., Meyer, A. N., Robertson, S. C. & Donoghue, D. J. (2000). Rotational coupling of the transmembrane and kinase domains of the Neu receptor tyrosine kinase. *Mol. Biol. Cell*, **11**, 3589–3599.
29. Chen, L. I., Webster, M. K., Meyer, A. N. & Donoghue, D. J. (1997). Transmembrane domain sequence requirements for activation of the p185c-neu receptor tyrosine kinase. *J. Cell Biol.* **137**, 619–631.
30. Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B. J., Dempsey, C. E. & Engelman, D. M. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J. Biol. Chem.* **267**, 7683–7689.
31. Sulistijo, E. S., Jaszewski, T. M. & MacKenzie, K. R. (2003). Sequence-specific dimerization of the transmembrane domain of the "BH3-only" protein BNIP3 in membranes and detergent. *J. Biol. Chem.* **278**, 51950–51956.
32. Arkin, I. T., Adams, P. D., MacKenzie, K. R., Lemmon, M. A., Brunger, A. T. & Engelman, D. M. (1994). Structural organization of the pentameric transmembrane alpha-helices of phospholamban, a cardiac ion channel. *EMBO J.* **13**, 4757–4764.
33. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T. & Engelman, D. M. (2000). Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nature Struct. Biol.* **7**, 154–160.
34. Xie, D., Shu, X. O., Deng, Z., Wen, W. Q., Creek, K. E., Dai, Q. *et al*. (2000). Population-based, case-control study of HER2 genetic polymorphism and breast cancer risk. *J. Natl Cancer. Inst.* **92**, 412–417.
35. Kuraoka, K., Matsumura, S., Hamai, Y., Nakachi, K., Imai, K., Matsusaki, K. *et al*. (2003). A single nucleotide polymorphism in the transmembrane domain coding region of HER-2 is associated with development and malignant phenotype of gastric cancer. *Int. J. Cancer*, **107**, 593–596.
36. Houliston, R. S., Hodges, R. S., Sharom, F. J. & Davis, J. H. (2004). Characterization of the proto-oncogenic and mutant forms of the transmembrane region of Neu in micelles. *J. Biol. Chem.* **279**, 24073–24080.
37. Fleming, K. G., Ackerman, A. L. & Engelman, D. M. (1997). The effect of point mutations on the free energy of transmembrane alpha-helix dimerization. *J. Mol. Biol.* **272**, 266–275.
38. Fisher, L. E., Engelman, D. M. & Sturgis, J. N. (1999). Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain. *J. Mol. Biol.* **293**, 639–651.
39. Doura, A. K., Kobus, F. J., Dubrovsky, L., Hibbard, E. & Fleming, K. G. (2004). Sequence context modulates the stability of a GxxxG-mediated transmembrane helix–helix dimer. *J. Mol. Biol.* **341**, 991–998.
40. Doura, A. K. & Fleming, K. G. (2004). Complex interactions at the helix–helix interface stabilize the glycophorin a transmembrane dimer. *J. Mol. Biol.* **343**, 1487–1497.
41. Fleming, K. G. (2002). Standardizing the free energy change of transmembrane helix–helix interactions. *J. Mol. Biol.* **323**, 563–571.
42. Kobus, F. J. & Fleming, K. G. (2005). The GxxxG-containing transmembrane domain of the CCK4 oncogene does not encode preferential self-interactions. *Biochemistry*, **44**, 1464–1470.
43. Yphantis, D. A. (1964). Equilibrium ultracentrifugation of dilute solutions. *Biochemistry*, **47**, 297–317.
44. Lemmon, M. A. & Schlessinger, J. (1994). Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem. Sci.* **19**, 459–463.
45. Schlessinger, J. (1988). Signal transduction by allosteric receptor oligomerization. *Trends Biochem. Sci.* **13**, 443–447.
46. Atkinson, A. C. (1980). A note on the generalized information criterion for choice of a model. *Biometrika*, **67**, 413–418.
47. Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36**, 575–588.
48. Pinkas-Kramarski, R., Soussan, L., Waterman, H., Levkowitz, G., Alroy, I., Klapper, L. *et al*. (1996). Diversification of Neu differentiation factor and epidermal growth factor signaling by combinatorial receptor interactions. *EMBO J.* **15**, 2452–2467.
49. Tzahar, E., Waterman, H., Chen, X., Levkowitz, G., Karunagaran, D., Lavi, S. *et al*. (1996). A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor. *Mol. Cell. Biol.* **16**, 5276–5287.
50. Langosch, D., Brosig, B., Kolmar, H. & Fritz, H. J. (1996). Dimerisation of the glycophorin A transmembrane segment in membranes probed with the ToxR transcription activator. *J. Mol. Biol.* **263**, 525–530.
51. Russ, W. P. & Engelman, D. M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl Acad. Sci. USA*, **96**, 863–868.
52. Fisher, L. E., Engelman, D. M. & Sturgis, J. N. (2003). Effect of detergents on the association of the glycophorin a transmembrane helix. *Biophys. J.* **85**, 3097–3105.
53. Lemmon, M. A., Treutlein, H. R., Adams, P. D., Brunger, A. T. & Engelman, D. M. (1994). A dimerization motif for transmembrane alpha-helices. *Nature Struct. Biol.* **1**, 157–163.
54. Melnyk, R. A., Partridge, A. W. & Deber, C. M. (2001).

Retention of native-like oligomerization states in transmembrane segment peptides: application to the *Escherichia coli* aspartate receptor. *Biochemistry*, **40**, 11106–11113.
55. Jones, D. H., Ball, E. H., Sharpe, S., Barber, K. R. & Grant, C. W. (2000). Expression and membrane assembly of a transmembrane region from Neu. *Biochemistry*, **39**, 1870–1878.
56. Li, S. C., Deber, C. M. & Shoelson, S. E. (1994). Glu-mediated dimerization of the transmembrane region of the oncogenic neu protein. In *Peptides: Chemistry, Structure, and Biology* (Hodges, R. S. & Smith, J. A., eds), pp. 757–759, ESCOM, Leiden, The Netherlands.
57. Rigby, A. C., Barber, K. R., Shaw, G. S. & Grant, C. W. (1996). Transmembrane region of the epidermal growth factor receptor: behavior and interactions *via* ^{2}H NMR. *Biochemistry*, **35**, 12591–12601.
58. Jones, D. H., Rigby, A. C., Barber, K. R. & Grant, C. W. (1997). Oligomerization of the EGF receptor transmembrane domain: a ^{2}H NMR study in lipid bilayers. *Biochemistry*, **36**, 12616–12624.
59. Morrow, M. R. & Grant, C. W. (2000). The EGF receptor transmembrane domain: peptide–peptide interactions in fluid bilayer membranes. *Biophys. J.* **79**, 2024–2032.
60. Sharpe, S., Grant, C. W., Barber, K. R., Giusti, J. & Morrow, M. R. (2001). Structural implications of a Val→Glu mutation in transmembrane peptides from the EGF receptor. *Biophys. J.* **81**, 3231–3239.
61. Sharpe, S., Barber, K. R. & Grant, C. W. (2002). Evidence of a tendency to self-association of the transmembrane domain of ErbB-2 in fluid phospholipid bilayers. *Biochemistry*, **41**, 2341–2352.
62. Hunter, T., Ling, N. & Cooper, J. A. (1984). Protein kinase C phosphorylation of the EGF receptor at a threonine residue close to the cytoplasmic face of the plasma membrane. *Nature*, **311**, 480–483.
63. Cochet, C., Gill, G. N., Meisenhelder, J., Cooper, J. A. & Hunter, T. (1984). C-kinase phosphorylates the epidermal growth factor receptor and reduces its epidermal growth factor-stimulated tyrosine protein kinase activity. *J. Biol. Chem.* **259**, 2553–2558.
64. Poppleton, H. M., Wiepz, G. J., Bertics, P. J. & Patel, T. B. (1999). Modulation of the protein tyrosine kinase activity and autophosphorylation of the epidermal growth factor receptor by its juxtamembrane region. *Arch. Biochem. Biophys.* **363**, 227–236.
65. Fleming, K. G., Ren, C. C., Doura, A. K., Eisley, M. E., Kobus, F. J. & Stanley, A. M. (2004). Thermodynamics of glycophorin A transmembrane helix dimerization in C14 betaine micelles. *Biophys. Chem.* **108**, 43–49.
66. Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. (2004). Quantification of helix–helix binding affinities in micelles and lipid bilayers. *Protein Sci.* **13**, 2600–2612.
67. Kwatra, M. M., Bigner, D. D. & Cohn, J. A. (1992). The ligand binding domain of the epidermal growth factor receptor is not required for receptor dimerization. *Biochim. Biophys. Acta*, **1134**, 178–181.
68. Moriki, T., Maruyama, H. & Maruyama, I. N. (2001). Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. *J. Mol. Biol.* **311**, 1011–1026.
69. Lemmon, M. A., Bu, Z., Ladbury, J. E., Zhou, M., Pinchasi, D., Lax, I. *et al*. (1997). Two EGF molecules contribute additively to stabilization of the EGFR dimer. *EMBO J.* **16**, 281–294.
70. Cho, H. S., Mason, K., Ramyar, K. X., Stanley, A. M., Gabelli, S. B., Denney, D. W. Jr & Leahy, D. J. (2003). Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature*, **421**, 756–760.
71. Garrett, T. P., McKern, N. M., Lou, M., Elleman, T. C., Adams, T. E., Lovrecz, G. O. *et al*. (2002). Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell*, **110**, 763–773.
72. Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J. H. *et al*. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell*, **110**, 775–787.
73. Franklin, M. C., Carey, K. D., Vajdos, F. F., Leahy, D. J., de Vos, A. M. & Sliwkowski, M. X. (2004). Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. *Cancer Cell*, **5**, 317–328.
74. Dawson, J. P., Melnyk, R. A., Deber, C. M. & Engelman, D. M. (2003). Sequence context strongly modulates association of polar residues in transmembrane helices. *J. Mol. Biol.* **331**, 255–262.
75. Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. (2001). Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl Acad. Sci. USA*, **98**, 2250–2255.
76. Gratkowski, H., Lear, J. D. & DeGrado, W. F. (2001). Polar side chains drive the association of model transmembrane peptides. *Proc. Natl Acad. Sci. USA*, **98**, 880–885.
77. Laue, T. M., Shah, B., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds), pp. 90–125, Royal Society of Chemistry, Cambridge.
78. Straume, M. & Johnson, M. L. (1992). Analysis of residuals: criteria for determining goodness-of-fit. *Methods Enzymol.* **210**, 87–105.
79. le Maire, M., Champeil, P. & Moller, J. V. (2000). Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta*, **1508**, 86–111.

Edited by G. von Heijne

(Received 3 December 2004; received in revised form 18 January 2005; accepted 23 January 2005)

CHAPTER 19

Analysis of Membrane Proteins Using Analytical Ultracentrifugation

KAREN G. FLEMING

1 Introduction

In recent years, analytical ultracentrifugation (AUC) has played an important role in the analysis of membrane proteins. Sedimentation velocity continues to be an extremely useful tool for assessing the homogeneity of membrane protein preparations in detergent micelle solutions; however, because of the multiple components present in a membrane protein:detergent micelle particle, the hydrodynamic interpretation of the sedimentation coefficient is not as straightforward as it is with soluble proteins. As in the past, sedimentation equilibrium continues to be used to determine the molecular weights of many membrane proteins when dissolved in detergent micelles. More recently, the method has played a major role in studies on the thermodynamics of subunit interactions between membrane proteins. The development of these protocols has spurned theories describing the dependence of a membrane protein oligomeric population on the amount of detergent as well as novel studies aimed at designing membrane proteins with desired oligomerization properties. The bulk of the contents of this chapter will therefore mainly focus on the issues related to measuring and interpreting the buoyant molecular weight of a membrane protein in complex with detergent.

1.1 General Considerations for Sedimentation Equilibrium Experiments to be Carried Out on Membrane Proteins

The general considerations for an experiment to be carried out on a membrane protein are similar to those applicable to a soluble protein. Since sedimentation equilibrium experiments can be very sensitive to small amounts of contaminating material, the best results will be obtained if the membrane protein preparation is $>95\%$ pure as judged by polyacrylamide gel electrophoresis. Due to their hydrophobic nature, purified integral membrane proteins in solution must be studied in buffers containing

a detergent micelle or lipid vesicle cosolvent, and the vast majority of AUC experiments are executed on membrane proteins dissolved in detergent micelles. Membrane proteins are typically analyzed using the absorbance optics because the interference optics would also contain a signal from the detergent, which would introduce uncertainty into monitoring the sedimentation of the particle of interest. Therefore, care must be taken to avoid buffer components that have significant absorbance at the detection wavelength of interest. Unless the membrane protein binds nucleic acids, this is not typically a concern for detection at 280 nm since most commonly used buffer components are transparent at 280 nm. However, some buffering components and reducing agents do have significant absorbance at 230 nm. Their use should be carefully monitored to avoid saturation of the absorbance optics system by the background buffer. Membrane proteins have been analyzed in the standard 2- and 6-sector cells using either quartz or sapphire windows. Sapphire windows are not always transparent at 230 nm, and the user should empirically determine the transparency for each set of windows.

The absorbance properties of the detergent micelles must also be considered. Most detergents are transparent at 280 nm; however Triton X-100 is a notable exception, and this detergent is not usually the optimal choice for this reason, although there have been studies carried out in low concentrations of Triton X-100.[1] Reduced Triton X-100 has been used more frequently because it has greater transparency at 280 nm. Neither of these detergents are a problem for membrane proteins containing chromophores, like heme, when the absorbance of the sedimenting particle can be monitored at a longer wavelength. Depending on the chemical structure, some detergents also have a modest absorbance at 230 nm. If the detergent concentration can be maintained such that this contribution is small (typically $<$0.4OD), then their use does not usually pose a significant experimental problem.

If a density-matching strategy (described below) is to be used, the detergent micelle environment should also be as homogeneous as possible. A mixture of detergent micelle types with different partial specific volumes could compromise the integrity of the density matching method, since a membrane protein may preferentially bind one type of detergent over the other. To obtain a homogeneous detergent environment, the membrane protein may be purified using that detergent. Alternatively, ion exchange chromatography can be used to exchange one detergent for another.[2–5] In some cases, if the critical micelle concentration is high enough, the detergent can be brought to equilibrium with the solution by dialysis.

1.2 Choice of Detergent Micelle Environment

The chemical diversity of detergent micelles types is large, and not surprisingly the effects of any particular detergent or lipidic environment upon any particular protein can vary widely.[6] Both the type and concentration can have profound effects on the integrity of a membrane protein and on the stability of any oligomers that the protein may form.[5–7] In addition to considerations of the effects of the detergent on the membrane protein functional properties, there may be experimental constraints on the choice of detergent micelles. These will depend on the type of analysis to be carried out. For a simple molecular weight determination of a membrane protein population

that is expected to be homogeneous, the available detergent choices are large since many of the strategies described below can be optimized for this purpose. However, in recent years, much progress has been made using sedimentation equilibrium to determine the thermodynamics of interactions between membrane proteins in detergent micelle environments. For these thermodynamic experiments, the detergent choices are limited to those that can be used with the density-matching approach. This is because the monomeric and oligomeric forms of the membrane protein may bind different amounts of detergent, in which case the density increment would be expected to be concentration-dependent.

2 Expressions for the Buoyant Molecular Weight

2.1 The Buoyant Molecular Weight Expressed in Terms of the Protein Moiety Alone

In the case of a membrane protein, the sedimenting particle will contain contributions from the sedimentation of the protein as well as from any bound detergent or lipid molecules. As is the case for soluble proteins, the sedimenting particle also contains a contribution from bound water, although it is negligible under most experimental conditions. By consideration of the thermodynamics of multicomponent systems it has been shown that the buoyant molecular weight of a membrane protein in complex with its hydrophobic solvent is measured in a sedimentation equilibrium experiment and equals[8]

$$M_P(1 - \phi'\rho) = \frac{2RT}{\omega^2}\left(\frac{d\ln c}{dr^2}\right) \quad (1)$$

where the buoyant molecular weight equals $M_P(1-\phi'\rho)$. In the above equation M_P is the molecular weight of only the protein portion of the sedimenting particle, which excludes the molecular weight of bound detergent, lipid and water, ϕ' the effective partial specific volume of the protein moiety in the sedimenting particle and takes into account the contributions of bound detergent, lipid and water; ρ the solvent density, R the gas constant, T the absolute temperature, ω the angular velocity (rad s^{-1}), c the concentration of the sedimenting particle, and r the radial distance from the center of rotation.

In a thermodynamically ideal solution, it has been shown that the buoyant molecular weight can be expressed in terms of its components[9,10]

$$M_P(1 - \phi'\rho) = M_P(1 - \overline{\nu}_p\rho) + \sum n_i M_i(1 - \overline{\nu}_i\rho) \quad (2)$$

where Equation (2) is the general equation, and M_i and $\overline{\nu}_i$ are the molecular weight and partial specific volumes (mL g^{-1}) of the ith component, and n_i is the number of molecules of any ith component bound to the protein. Since most analysis of membrane proteins is carried out on purified proteins dispersed in a detergent micelle solution, we can explicitly write the buoyant molecular weight as

$$M_P(1 - \phi'\rho) = M_P(1 - \overline{\nu}_P\rho) + n_{Det}M_{Det}(1 - \overline{\nu}_{Det}\rho) + n_{H_2O}M_{H_2O}(1 - \overline{\nu}_{H_2O}\rho) \quad (3)$$

where the subscripts Det and H_2O indicate the contributions from bound detergent and water molecules, respectively. This equation is also written in the form

$$M_P(1 - \phi'\rho) = M_P[(1 - \overline{\nu}_P\rho) + \delta_{Det}(1 - \overline{\nu}_{Det}\rho) + \delta_{H_2O}(1 - \overline{\nu}_{H_2O}\rho)] \quad (4)$$

where δ_i represents the amount of the ith component bound in grams per gram of protein. In Equations (3) and (4) it is assumed that the contribution from lipid is negligible, however in some cases, tightly bound lipid molecules will co-purify with a membrane protein. If the number of lipids is large, then an additional term accounting for the bound lipid must also be included in Equations (3) and (4); however small amounts of lipid can usually be ignored because the partial specific volumes for lipids are generally near unity.[10] For clarity in this initial discussion, we will assume that negligible amounts of lipid are present, and examples where this is not the case will be discussed later. In addition, it is important to note that the $\overline{\nu}_{Det}$ term in Equations (3) and (4) represents the partial specific volume of the detergent when it is bound to the protein, which is one of the fundamental assumptions in this analysis. Tanford and Reynolds[10] has argued that the partial specific volume of detergents above the critical micelle concentration should be used because this represents the $\overline{\nu}_{Det}$ when detergent is in a self-associated state, as it would be when associated with a membrane protein.

From Equations (3) and (4), it can be seen that the experimentally measured buoyant molecular weight contains information about both the membrane protein molecular weight and the molecular weight of any bound detergent. Depending on the scientific question, the experimental conditions can be optimized for determining the value of one or both of these parameters.

2.2 The Buoyant Molecular Weight Expressed in Terms of the Protein–Detergent Complex

An alternative theoretical expression for the buoyant molecular weight of the sedimenting particle was first used by Hersh and Schachman[9] in 1958 and assumes that the partial specific volume of the complex is approximated by the sum of the volumes of the protein and detergent. This theoretical formalism expresses the quantity measured in a sedimentation experiment more directly in terms of the protein–detergent particle:

$$M_C(1 - \overline{\nu}_C\rho) = \frac{2RT}{\omega^2}\left(\frac{\mathrm{d}\ln c}{\mathrm{d}r^2}\right) \quad (5)$$

where here the buoyant molecular weight is $M_C(1-\overline{\nu}_C\rho)$, with M_C being the molecular weight of the entire protein detergent particle which equals $M_P(1+\delta_{Det})$ and $\overline{\nu}_C$ equals the partial specific volume of the entire complex and can be written as

$$\overline{\nu}_C = \frac{\overline{\nu}_P + \delta_{Det}\overline{\nu}_{Det}}{1 + \delta_{Det}} \quad (6)$$

Additional terms can be added if a contribution from lipid needs to be taken into account. It is important to note that the buoyant molecular weights in Equations (1)

and (5) must be equal, which emphasizes the point that $\phi' \neq \overline{v}_C$ because $M_P \neq M_C$. Expression (5) is included here because there is sometimes confusion in the literature about the definition of the buoyant molecular weight for a membrane protein in detergent solutions. In practice, the theoretical formalism in Equation (5) is not used as widely and will not be discussed extensively in this chapter. Tanford and Reynolds[10] has pointed out that, using the expression in Equation (1), the quantity M_P may be determined with greater accuracy than the quantity M_C from Equation (5) because $\overline{v}_C$ is not easily measured. In contrast, as will be discussed below, the quantity $(1-\phi'\rho)$ can be directly determined by density measurements.

3 Sedimentation Equilibrium Methods to Determine Protein Molecular Weights in Protein–Detergent Complexes

3.1 Determination of Membrane Protein Molecular Weights by Classical Density Matching

The molecular weight of only the protein portion of the buoyant molecular complex can be determined by adjusting the experimental conditions such that the net contribution of the detergent becomes negligible. This method is cartooned in Figure 1 and has been referred to as "density matching" or the condition of rendering the detergent gravitational transparent.[11] Density matching can be accomplished by modifying the solvent density such that it becomes equal to the effective density of the bound detergent, *e.g.* $\rho = 1/\overline{v}_{Det}$. When this is the case, the term $(1-\overline{v}_{Det}\rho)$ approaches zero, and the contribution from the detergent term in Equation (3), $n_{Det}M_{Det}(1-\overline{v}_{Det}\rho)$ becomes negligible regardless of the number of bound detergent molecules.

Thermodynamically, the most favorable experimental condition can be obtained when the density of the solvent can be modulated using a mixture of light and heavy water. This experimental condition produces the most minimal perturbation to the contribution of water in the buoyant molecular weight of the sedimenting species. When the density matching of detergent is carried out using mixtures of water, the average effective density of any bound water is assumed to be equal to the density of the solvent. In this case, the water contribution term from Equation (3), $\overline{v}_{H_2O}M_{H_2O}(1-\overline{v}_{H_2O}\rho)$, will also be essentially equal to zero. The result of such a density matching strategy is that Equation (3) can be reduced to

$$M_P(1 - \phi'\rho) = M_P(1 - \overline{v}_P\rho) \tag{7}$$

Assuming that the partial specific volume of the protein can be calculated from the amino acid composition using standard procedures,[12] the buoyant molecular weight of the sedimenting complex can be directly interpreted in terms of the molecular weight of the protein portion alone.

To match the effective density of the detergent micelles, the most commonly used water mixture is H_2O/D_2O. Experiments at higher solvent densities can be carried out using $D_2{}^{18}O$, although this has been used less extensively due to the expense of

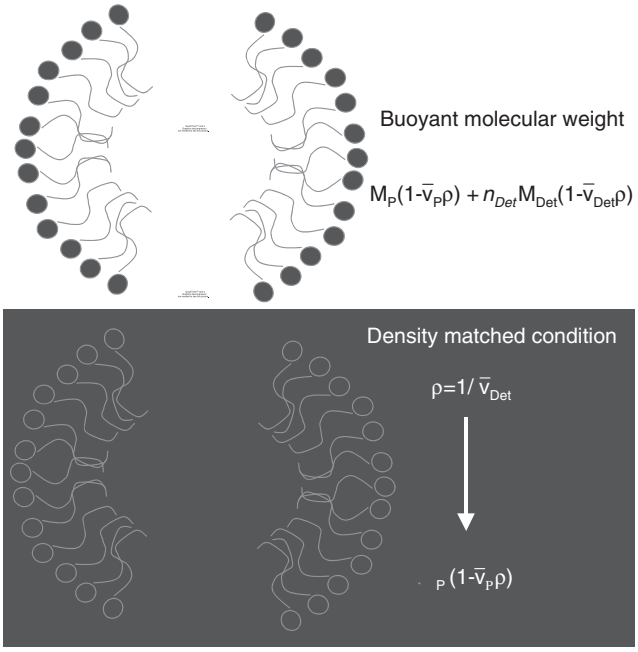


Figure 1 *Cartoon of density matching strategy. Away from the density match point, both the protein and the detergent contribute to the sedimentation of the particle at equilibrium. If the solvent density is adjusted such that it equals the effective density of the bound detergent as inlower panel of the figure, the bound detergent becomes essentially invisible to the gravitational field, and the sedimentation of the particle is determined by the protein moiety alone. This cartoon represents the classic density matching strategy where the density is matched with heavy water, and so no distinction is made for bound water, which is density-matched in both cases in this figure*

the $D_2{}^{18}O$ compound.[13] Using deuterium water mixtures, there is a small mass correction that must be incorporated into the analysis due to the exchange of deuterium for hydrogen on exchangeable protons.[13,14] The density of heavy water is limited to 1.10392 and 1.21465 g mL^{-1} for 100% D_2O and 100% $D_2{}^{18}O$, respectively (at 20 °C).[12] For this reason, the use of the density-matching strategy places limitations on the detergents that can be employed in sedimentation equilibrium experiments with this method. To match the effective density, the partial specific volumes of the detergent must be greater than 0.9058 mL g^{-1} to match with D_2O and greater than 0.8232 g mL^{-1} to match with $D_2{}^{18}O$. It is also important to keep in mind that detergents and lipids with partial specific volumes greater than unity will have effective densities that are less than the density of normal water. This will cause a negative contribution to the buoyant molecular weight and be manifested as a tendency to float in the ultracentrifugation experiment.[15] Due to these density considerations, the detergents most commonly used with the classic density-matching strategy are C_8E_4, C_8E_5, $C_{12}E_8$, C_{14} sulfobetaine, and dodecylphosphocholine (DPC).

The first step in a density-matching experiment is to establish the conditions under which the detergent micelles are gravitationally transparent. A good starting point is

the consideration of the partial specific volume for the detergent above the critical micelle concentration, and the partial specific volumes for many detergents are known.[10,16–18] When the partial specific volume is not known, it can be determined by at least two methods. Using a high-precision densimeter, direct density measurements can be used to measure the partial specific volume of a detergent micelle solution in a buffer of interest.[17] Alternatively, the analytical ultracentrifuge can be used to measure the buoyant molecular weight of the detergent micelles alone in a series of solvents whose densities differ. Since many partial specific volumes have been tabulated, this approach is typically used to confirm the percentage of D_2O required for matching the effective density of the detergent micelle in the context of the other buffering components. At the density-match point, the detergent micelles will be uniformly dispersed within the cell. At slightly lower or higher solution densities, the distributions are usually linear with positive and negative slopes, respectively. For a rigorous analysis, the value for the detergent micelle buoyant molecular weight can be determined using Equation (1), and the detergent micelles are density matched when this buoyant molecular weight equals zero. It is optimal to use interference optics to detect the detergent micelle radial distribution when the detergent micelles are transparent at accessible wavelengths. Although the micelles can be visualized using a hydrophobic dye that partitions into them, we have found that even a small amount of dye can alter the micelle density.

3.2 Examples of Density Matching

Several groups have used the density-matching strategy to determine molecular weights of membrane proteins. Classic work by Reynolds and Tanford first demonstrated the usefulness of the density-matching strategy.[10,18,19] Since then, many studies have been carried out with increasing frequency in the past few years. Several groups have taken advantage of the properties of the C_8E_4 or C_8E_5 detergents, which are neutrally buoyant and require no D_2O for density matching. The molecular weights of membrane proteins with diverse folds including monotopic and polytopic α-helical membrane proteins as well as transmembrane β-barrel proteins have been analyzed using these neutrally buoyant detergent micelles.[20–26] $C_{12}E_8$ micelles have also been used in the analysis of membrane proteins by sedimentation equilibrium. In one study, its effective density was matched with 18% D_2O.[27] In a different study the relative mass contribution due to bound $C_{12}E_8$ was assumed to be small and no D_2O was added.[28] More recently, two groups have used C_{14} sulfobetaine micelles matched with 13% D_2O or DPC micelles matched with 50% D_2O to determine the protein molecular weights of both natural and designed membrane protein sequences.[5,29–36] The density-matching strategy has especially found utility in the thermodynamic analysis of membrane protein interactions.

3.3 When Heavy Water is not Heavy Enough

The effective densities of some of the most widely used types of detergent micelles exceed that of D_2O and therefore cannot be matched using heavy water. This class includes dodecylmaltoside (DDM) and β-D-octylglucoside (βOG), which are both

commonly used to dissolve membrane proteins for functional and structural studies.[6] When it is optimal for the membrane protein of interest to be solubilized in these types of detergents or other detergents with partial specific volumes $\lesssim$0.9 mL g^{-1}, alternative strategies to density matching must be used to disentangle the relative contributions of the protein and detergent contributions to the buoyant molecular weight. Two strategies for the analysis of membrane proteins with density agents will be discussed.

3.3.1 Extrapolation to the Match Point

The first method involves determining the buoyant molecular weight in several different mixtures of heavy water followed by extrapolation to the density at which the detergent micelles are gravitationally transparent.[19] If the protein–detergent complex is not altered by changing the H_2O/D_2O (or $D_2{}^{18}O$) ratio, the buoyant molecular weight of the sedimenting complex will be a linear function of solvent density. A line can be fitted and at the match point for the effective density of the detergent, the value of the buoyant molecular weight can be interpreted in terms of the protein portion alone using Equation (3) since the two right terms in Equation (3) are negligible at this match point. A good example of this method is shown in Figure 2. Suarez *et al.*[13] used this extrapolation approach to measure the partial specific volume of DDM micelles as well as the molecular weights of both heart and rat liver cytochrome oxidase dissolved in the same detergent. As can be seen in Figure 2, the extrapolated line for the buoyant molecular weight of DDM micelles crosses zero at a solvent density of 1.22 g mL^{-1}. At this same point, which is indicated by the dotted vertical line, the buoyant molecular weights of the cytochrome oxidase proteins are still positive, reflecting the sedimentation of the protein portion. At this extrapolated match point the molecular weight of the protein portion alone can be calculated using Equation (4) by knowing the protein partial specific volume, which can be calculated from the amino acid composition.[12, 37] This extrapolation method has the disadvantage of requiring an extrapolation to a match point that is not experimentally accessible, however it can be accurate if sufficient data are collected to describe the line and if the extrapolation is not too long. In fact, Reynolds and mcCaslin[18] has even used this method to estimate the molecular weight of bovine rhodopsin in sodium cholate micelles, whose partial specific volume is nearly that of proteins.[18]

3.3.2 Determination of Membrane Protein Molecular Weights by Density Matching the Hydrated Detergent Micelle with Density Agents

A second experimental approach recently employed is the use of molecules to increase the solution density, which will be referred to as densifier molecules. This experimental approach is most successfully implemented when moderate amounts of the density agents can successfully match the sedimentation of the micelles. Nycodenz, sucrose, and glycerol have all been employed in several different studies.[11,38,39] Since the addition of these densifier compounds can lead to preferential binding or exclusion of water, the effective density of the detergent usually contains a contribution from bound water. For this reason, this method is sometimes referred to as the density-matching of

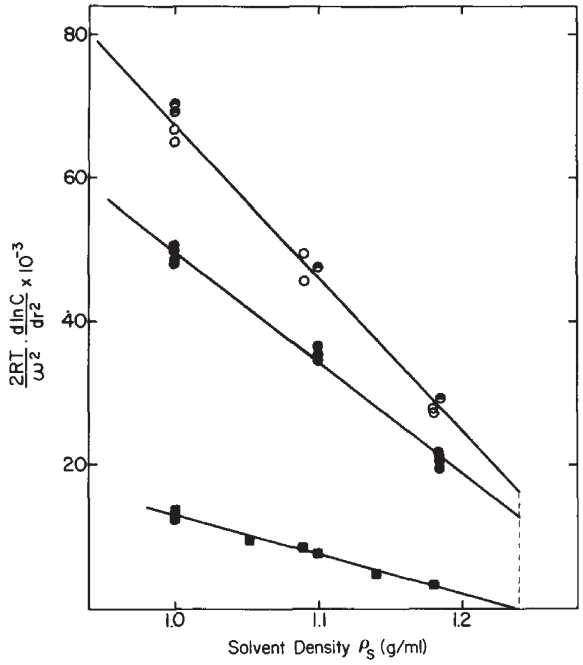


Figure 2 *An example of the method of extrapolation to the density match point using heavy water. The buoyant molecular weights for beef heart cytochrome oxidase (open circles), rat liver cytochrome oxidase (closed circles), and DDM micelles (squares) were determined as a function of solvent density. The dotted vertical line indicates the X-intercept for the linear fit to the sedimentation of DDM micelles; at which point the DDM micelles are density-matched. The increased mass for the two protein samples at this solvent density reflects the sedimentation due to the protein portion of the sedimenting complex for those oxidase proteins. Using the protein molecular weight and partial specific volume for DDM determined at the match point for DDM, the amount of detergent bound to each protein can be calculated from the buoyant molecular weights at the other solvent densities using Equation (3). Reproduced with permission from Suarez et al.*[13]

the hydrated detergent micelles. Like the classic density-matching approach carried out with H_2O/D_2O mixtures, the first step with this experimental approach is to establish the density matching condition. As shown in Figure 3, this is accomplished by carrying out a sedimentation equilibrium experiment on detergent micelles alone as a function of density, where the density is modulated by the concentration of the density agent. Like the original method, the effective density of the hydrated detergent micelle is matched when the buoyant molecular weight equals zero. It has been found that the effective density of the hydrated detergent micelle varies significantly as a function of density agent, and it is generally lower than the effective density in H_2O/D_2O solutions, demonstrating the effect of the bound water.

In principle, the protein moiety may also experience preferential hydration or exclusion of water, and this will be reflected in the effective partial specific volume term for the protein. In practice, this contribution has been observed to be small as long as mod-

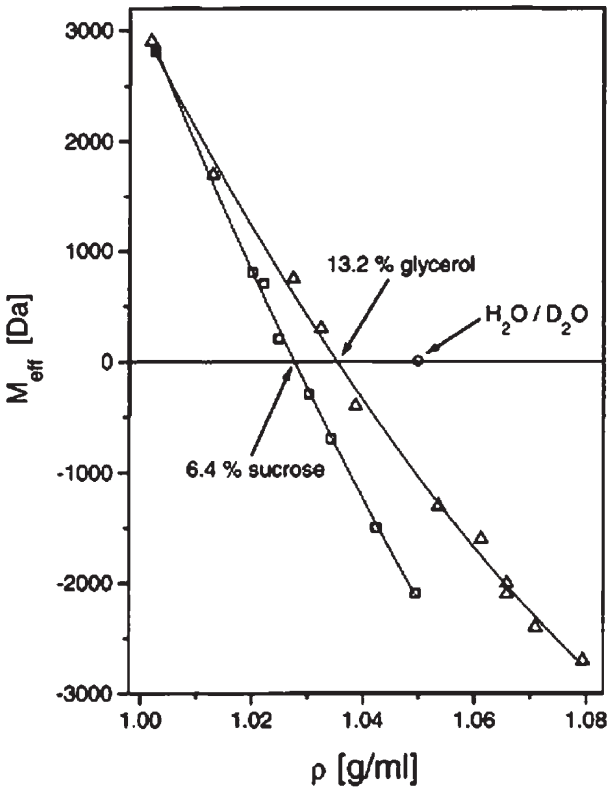


Figure 3 *An example of establishing the density-matching conditions for $C_{12}E_9$ micelles using sucrose or glycerol. The equilibrium sedimentation of 0.3% (w/v) $C_{12}E_9$ micelles was measured using interference optics in the presence of varying amounts of sucrose (open squares), glycerol (open triangles) or DwO (open circle). At equilibrium, the effective mass of the hydrated $C_{12}E_9$ micelles was determined as*

$$M_{\text{eff}} = M_{\text{Det}}(1-\overline{v}_{\text{Det}}\rho) = \frac{2RT}{\omega^2}\left(\frac{\mathrm{d}\ln c}{\mathrm{d}r^2}\right)$$

Reproduced with permission from Mayer et al.[38]

erate levels of the density agent are used. Lustig *et al.*[11] used an empirical approach to estimate the contribution of water to the partial specific volumes of six different membrane proteins and obtained protein molecular weights that agreed within ±15% of the known values in both DDM and βOG micelles. Lebowitz and co-workers used reduced Triton X-100 micelles density matched with sucrose to demonstrate the oligomeric state for simian immunodeficiency viral envelop protein particles.[39] These authors assumed a negligible contribution of bound water to the protein moiety and still were able to demonstrate a trimeric stoichiometry for the viral envelop protein particles. Mayer *et al.*[38] showed in $C_{12}E_9$ and reduced Triton X-100 that the partial specific volumes for the cytochrome *c* oxidase protein depended only slightly on the nature of the density agent and were similar to the value obtained in H2O/D_2O mixtures. However,

this same group found that experiments in DDM micelles were difficult to interpret when matched with either sucrose or glycerol alone in water.[38] Even in a subsequent study[40] the combination of sucrose with D_2O, which reduced the sucrose requirement, still required some assumptions about the contribution of hydration to the protein moiety in the presence of DDM micelles since the use of the partial specific volume calculated from the composition using the values of Cohn and Edsall[37] returned a protein molecular weight differing from the known value by a significant amount. In sum, if the hydrated detergent micelles can be matched by moderate amounts of density agent, this density-matching approach appears a promising method for molecular weight determinations in a wide variety of detergent micelle types.

Mayer and co-workers have pointed out that many of these densifier agents are known to stabilize proteins, and this may offer an additional advantage due to the relatively long periods of time required for sedimentation equilibrium experiments.[38] On the other hand, significant amounts of densifier agent can increase the solvent viscosity, which can increase the time to equilibrium. Thus, the advantages of this approach may depend on the molecular properties of the membrane protein in a particular detergent micelle environment.

3.4 Determination of Membrane Protein Molecular Weights by the Analysis of Density Increment

An alternative to density matching is a consideration of the buoyant molecular weight of the sedimenting particle directly in terms of the density increment. As defined by Casassa and Eisenberg,[8] in a thermodynamic consideration of multicomponent systems, the density increment for the macromolecule at constant chemical potential of water and all diffusible solutes (including detergent) equals

$$\left(\frac{\partial \rho}{\partial c_2}\right)_{\mu} = (1 - \phi' \rho_0) \tag{8}$$

where c_2 is the weight concentration of the protein alone, and the change in the solution density as a function of changing the protein concentration gives the density increment. This term $(1-\phi'\rho_0)$ is identical to the buoyancy factor given in Equation (1) where ϕ' is defined as before. When multiplied by the molecular weight of the protein moiety alone as in Equation (1), this quantity returns the buoyant molecular weight of the entire sedimenting complex. The advantage of analysis using the density increment method is that the quantity $(\partial\rho/\partial_{C2})_{\mu}$ can be directly and independently measured on a protein solution at dialysis equilibrium with all other components using a high-precision densimeter. As little as two density measurements on solutions in the presence and absence of protein has been used to calculate the subsequent slope, which is applied to the sedimentation results to calculate the accurate of the molecular weights membrane protein. As long as accurate and precise density measurements can be made, this method greatly expands the repertoire of detergent micelles that can be used in a sedimentation equilibrium experiment since no density-matching considerations are necessary. In addition, this approach works well for membrane proteins dispersed in mixtures of different types of detergent micelles as well as detergent/lipid mixtures,

since there is no uncertainty due to preferential binding of one of the hydrophobic components.

3.5 Examples of Use of the Density Increment

Butler and co-workers have used this approach to determine the average molecular weights of membrane proteins dispersed in a wide variety of detergent micelles. The versatility of the method has been demonstrated for membrane proteins samples having effective partial specific volumes ranging from –0.14 to 0.734 mL g^{-1}.[41,42] The density increment method overcomes many of the difficulties described above associated with using the class of micelles containing maltoside head groups, as several membrane proteins have been measured in either DDM or decylmaltoside (DM).[43] The stoichiometry of the bovine major intrinsic protein was shown to be tetrameric in DM, one of the few detergents that stabilized the complex,[44] and the versatility of the density increment approach is demonstrated by recent studies aimed at determining the oligomeric state of the EmrE multidrug transporter solubilized in DDM detergent.[41] Two forms of the protein were prepared. A "high molecular weight" form that was purified using minimal detergent concentrations as well as a "lower molecular weight" form purified using higher concentrations of DDM. Sedimentation equilibrium demonstrated that both forms of the EmrE were in fact dimeric. In complementary work that determined the amount of DDM bound to EmrE, it was found that the "high molecular weight" form of EmrE was not a different oligomeric state, but in fact had ~10-fold more lipid bound. This additional lipid increased the apparent size as measured by gel filtration and sedimentation velocity and obscured true molecular weight analysis using those methods.

3.6 Thermodynamics of Membrane Protein Association

The origins of the specificity and free energy of interaction between membrane proteins are not well understood. Probably the largest impact that sedimentation equilibrium has had on membrane proteins in recent years has been due to its use in experimental measurements in this area of research. While it had been widely appreciated that the properties of micelles influenced the functional and oligomeric states of membrane proteins, this knowledge was derived from empirical observations of the effects of particular detergents on specific proteins.[6] Using sedimentation equilibrium, Fleming and co-workers have shown that the free energy of association between membrane proteins can be measured in detergent micelle solutions as long as the kinetics of rearrangement are fast relative to the time for sedimentation equilibrium.[2,4,27]

The experimental advances lead to several theories describing the concentration dependence of interactions in micelles.[4,5,28,45] Under the assumption that a self-association reaction for an integral membrane protein will take place within the micellar environment, it has been shown in two different detergents that the apparent equilibrium constant for a monomer–dimer reaction varies in a systematic way as a function of detergent concentration.[4] This apparent equilibrium constant is derived from the sedimentation equilibrium data using the protein concentration in bulk molar units. As can be seen in Figure 4, the free energy of association will become less favorable if the

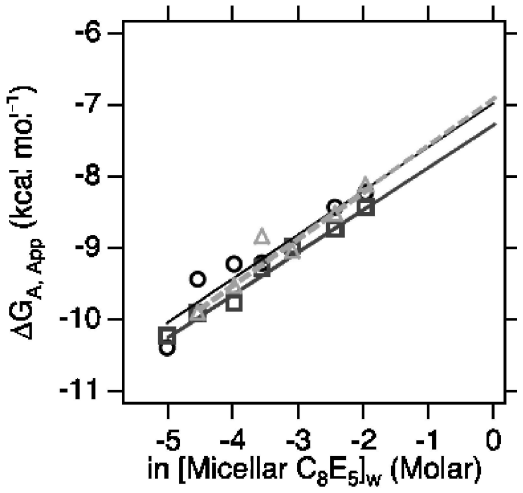


Figure 4 *The apparent free energy of association for a membrane protein monomer–dimer reaction decreases as the size of the micelle phase increases. The apparent free energy of association is plotted as a function of the logarithm of the micellar detergent concentration expressed on the aqueous scale. Under the ideal dilute assumption, the data should vary linearly with a slope of RT. The linear regressions of three independent experiments an average slope of 0.62 (RT = 0.59 at 25 °C) Experiments in C_8E_5 micelles were carried out on the glycophorin A transmembrane helix expressed as a fusion protein with Staphylococcal nuclease. Reproduced with permission from Fleming, 2002*[4]

size of the micellar phase increases because the membrane protein population experiences an increase in volume of the reaction phase. The large number of data points available in a sedimentation equilibrium experiment coupled with modern computational power allowing global fits of different initial starting conditions are two advantages of using the analytical ultracentrifuge for these studies over alternative methods.

The availability of protocols to quantify the free energy of association between integral membrane proteins in detergent solutions has catalyzed a new area of research in the membrane protein field. Several mutagenesis studies on model systems as well as natural membrane proteins have begun to address the physical principles underlying oligomeric stability of membrane proteins.[2,25,26,36] These principles are being tested by membrane protein design efforts whose success are evaluated in large part by sedimentation equilibrium analysis of the designed sequences.[29–33,46]

3.7 Determination of the Amount of Detergent Bound in a Membrane Protein Complex

In the characterization of a membrane protein complex, the amount of detergent bound to a membrane protein may be of interest for a particular scientific question. Implicit in many of the discussions above is the fact that sedimentation equilibrium experiments carried out at densities away from the detergent match point contain a measurable contribution from the bound detergent. Once the molecular weight of the

protein moiety has been determined, Equation (3) can be used to calculate the quantity n_{Det} as long as the contribution from hydration is small.[13] More recently, modern computers have been used to globally fit the primary sedimentation equilibrium data collected at different D_2O concentrations directly for this parameter.[47]

4 Sedimentation Velocity Analysis

While most of the recent advances in experimental methods for analysis of membrane proteins have come in the area of sedimentation equilibrium, sedimentation velocity still remains an extremely useful tool. Since finding optimal conditions for solubilization of a membrane protein can be an arduous task, sedimentation velocity continues to be a fairly quick experimental approach for evaluating the homogeneity of a sample.[11,48,49] A good example of this approach is shown in Figure 5, where Musatov *et al.* used the van Holde–Weischet analysis method to evaluate the sedimentation coefficient distribution for cytochrome *bo3* as a function of the concentration of Triton X-100. The sedimentation coefficient distribution showed that the protein aggregated when the detergent concentration was low (open squares). In contrast, a homogeneous sedimentation coefficient distribution was found at detergent/protein ratios >2:1 (w/w). Subsequent sedimentation equilibrium analysis demonstrated that the experimental conditions producing the homogeneous population represented a monomeric form of the cytochrome *bo3* membrane protein.

A similar question of homogeneity was of interest in a study by MacPhee *et al.*, who used sedimentation velocity to characterize the interactions of lipoprotein lipase

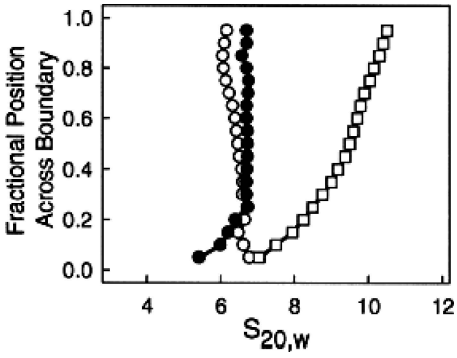


Figure 5 *An example of sedimentation velocity analysis to determine homogeneity in a membrane protein preparation. Sedimentation velocity data collected at 27 000 rpm were analyzed using the van Holde–Weischet method. Each line of symbols represents the distribution of sedimentation coefficients across the boundary. Cytochrome bo3 (0.5 mg protein mL^{-1}) was solubilized in 50 mM potassium phosphate (pH 8.1) containing (a) 1.0 mg mL^{-1} Triton X-100 (open squares), (b) 5 mg mL^{-1} Triton X-100 (filled circles) or (c) 10 mg mL^{-1} Triton X-100 (open circles). The near identity of sedimentation coefficients across the boundary as in samples b and c indicate sample homogeneity. Subsequent sedimentation equilibrium analysis demonstrated a monomeric population for cytochrome bo3 under conditions where the sedimentation coefficient indicated homogeneity. Reproduced with permission from Musatof et al.*[49]

with homogeneous lipid emulsions.[15] An interesting aspect of this work is that the density of the lipid emulsions is less than that of water. Therefore, the sedimentation velocity experiment was used to evaluate the flotation gradients formed upon application of the gravitational potential. This resulted in the quantification of complexes in a specialized version of the time derivative software[50] by defining a flotation coefficient, which corresponds to the flotation equivalent of a sedimentation coefficient. Flotation equilibrium studies were also carried out in this study.

5 The Future for Membrane Proteins and the Analytical Ultracentrifuge

As the development of expression systems to generate purified membrane proteins continues, an increasing number of membrane protein samples will be available for molecular weight and thermodynamic analysis using sedimentation equilibrium. The increase in availability of modern ultracentrifuges will make the method more accessible, and sedimentation velocity will continue to find utility as a method for screening membrane protein preparations for homogeneity. The exponential increase in high-resolution structural studies on membrane proteins,[51] most of which are carried out in detergent micelles, will stimulate solution studies on this class of proteins in order to test functional questions arising from the structures. The ultracentrifuge will continue to be a biophysical tool of choice for answering many of them.

Acknowledgments

The author gratefully acknowledges the General Medical Sciences Institute of the US National Institutes of Health, the US National Science Foundation, and the US Department of Defense Breast Cancer Research Fund for support of past and current research on membrane proteins.

References

1. H. Hackenberg and M. Klingenberg, *Biochemistry*,1980, **19**, 548–555.
2. K. G. Fleming and D. M. Engelman, *Proc. Natl. Acad. Sci., USA*, 2001, **98**, 14340–14344.
3. K. G. Fleming, A. L. Ackerman and D. M. Engelman, *J. Mol. Biol.,* 1997, **272**, 266–275.
4. K. G. Fleming, *J. Mol. Biol.,* 2002, **323**, 563–571.
5. K. G. Fleming, C. C. Ren, A. K. Doura, F. J. Kobus, M. E. Eisley and A. M. Stanley, *Biophys. Chem*., 2004, **108**, 43–49.
6. M. le Maire, P. Champeil and J. V. Møller, *Biochim. Biophys. Acta*, 2000, **1508**, 86–111.
7. M. le Maire, K. Kwee, J. P. Anderson and J. V. Møller, *Eur. J. Biochem.,* 1983, **129**, 525–532.
8. E. F. Casassa, and H. Eisenberg, *Adv. Prot. Chem*., 1964, **19**, 287–395.
9. R. T. Hersh and H. K. Schachman, *Virology,* 1958, **6**, 234–243.
10. C. Tanford and J. A. Reynolds, *Biochim. Biophys. Acta*, 1976, **457**, 133–170.
11. A. Lustig, A. Engel, G. Tsiotis, E. M. Landau and W. Baschong, *Biochim. Biophys. Acta,* 2000, **1464**, 199–206.
12. T. M. Laue, B. Shah, T. M. Ridgeway and S. L. Pelletier, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, S. E. Harding, A. J. Rowe and J. C. Horton, (eds), Royal Society of Chemistry, Cambridge, UK, 1992, 90–125.

13. M. D. Suarez, A. Revzin, R. Narlock, E. S. Kempner, D. A. Thompson and S. Ferguson-Miller, *J. Biol., Chem*., 1984, **259**, 13791–13799.
14. S. J. Edelstein, and H. K. Schachman, *Method. Enzymol*., 1973, **27**, 82–99.
15. C. E. MacPhee, R. Y. Chan, W. H. Sawyer, W. F. Stafford and G. J. Howlett, *J. Lipid Res*., 1997, **38**, 1649–1659.
16. H. Durshlag, in *Thermodynamic Data for Biochemistry and Biotechnology,* H.-J., Hinz, (ed), Springer, Berlin, 1986, 45–128.
17. C. Tanford, Y. Nozaki, J. A. Reynolds and S. Makino, *Biochemistry,* 1974, **13**, 2369–2376.
18. J. A. Reynolds and D. R. McCaslin, *Method. Enzymol*., 1985, **117**, 41–53.
19. J. A. Reynolds and C. Tanford, *Proc. Natl. Acad. Sci., USA*, 1976, **73**, 4467–4470.
20. B. Ludwig, M. Grabo, I. Gregor, A. Lustig, M. Regenass and J. P. Rosenbusch, *J. Biol. Chem*., 1982, **257**, 5576–5578.
21. K. P. Locher and J. P. Rosenbusch, *Eur. J. Biochem.,* 1997, **247**, 770–775.
22. S. Hellstern, S. Pegoraro, C. B. Karim, A. Lustig, D. D. Thomas, L. Moroder and J. Engel, *J. Biol., Chem*., 2001, **276**, 30845–30852.
23. A. M. Stanley and K. G. Fleming, *J. Mol. Biol.,* 2005, in press. **AQ1**
24. F. J. Kobus and K. G. Fleming, *Biochemistry*, 2005, in press. **AQ2**
25. A. K., Doura, F. J. Kobus, L. Dubrovsky, E. Hibbard and K.G. Fleming, *J. Mol. Biol*., 2004, **341**, 991–998.
26. A. K. Doura and K. G. Fleming, *J. Mol. Biol*., 2004, **343**, 1498–1497.
27. K. G. Fleming, *Method. Enzymol*., 2000, **323**, 63–77.
28. D. Josse, C. Ebel, D. Stroebel, A. Fontaine, F. Borges, A. Echalier, D. Baud, F. Renault, M. Le Maire, E. Chabrieres and P. Masson, *J. Biol. Chem*., 2002, **277**, 33386–33397.
29. C. Choma, H. Gratkowski, J. D. Lear and W. F. DeGrado, *Nat. Struct. Biol*. 2000, **7**, 161–166.
30. J. D. Lear, H. Gratkowski, L. Adamian, J. Liang and W. F. DeGrado, *Biochemistry*, 2003, **42**, 6400–6407.
31. J. D. Lear, A. L. Stouffer, H. Gratkowski, V. Nanda and W. F. Degrado, *Biophys. J*., 2004, **87**, 3421–3429.
32. H. Gratkowski, J. D. Lear, and W. F. DeGrado *Proc. Natl. Acad. Sci., USA*, 2001, **98**, 880–885.
33. H. Gratkowski, Q. H. Dai, A. J. Wand, W. F. DeGrado and J. D. Lear, *Biophys. J.* 2002, **83**, 1613–1619.
34. D. Salom, B. R. Hill, J. D. Lear and W. F. DeGrado, *Biochemistry*, 2000, **39**, 14160–14170.
35. R. Li, C. R. Babu, J. D. Lear, A. J. Wand, J. S. Bennett and W. F. DeGrado, *Proc. Natl. Acad. Sci., USA*, 2001, **98**, 12462–12467.
36. K. P. Howard, J. D. Lear and W. F. DeGrado, *Proc. Natl. Acad. Sci., USA*, 2002, **99**, 8568–8572.
37. E. J. Cohn and J. T. Edsall, in *Proteins, Amino Acids and Peptides*, E. J. Cohn and J. T. Edsall, (eds), Reinhold Publishing Corporation, New York, 1943, 370–381.
38. G. Mayer, B. Ludwig, H.-W. Muller, J. A. van den Broek, R. H. Friesen and D. Schubert, *Prog. Coll. Polym. Sci*., 1999, **113**, 176–181.
39. R. J. Center, P. Schuck, R. D. Leapman, L. O. Arthur, P. L. Earl, B. Moss and J. Lebowitz, *Proc. Natl. Acad. Sci., USA,* 2001, **98**, 14877–14882.
40. G. Mayer, O. Anderka, B. Ludwig and D. Schubert, *Prog. Coll. Polym. Sci.,* 2002, **119**, 77–83.
41. P. J. Butler, I. Ubarretxena-Belandia, T. Warne and C. G. Tate, *J. Mol. Biol.,* 2004, **340**, 797–808.
42. P. J. Butler and W. Kuhlbrandt, *Proc. Natl. Acad. Sci., USA,* 1988, **85**, 3797–3801.

43. E. H. Heuberger, L. M. Veenhoff, R. H. Duurkens, R. H. Friesen and B. Poolman, *J. Mol. Biol.,* 2002, **317**, 591–600.
44. N. Konig, G. A. Zampighi and P.J. Butler, *J. Mol. Biol.,* 1997, **265**, 590–602.
45. L. E., Fisher, D. M. Engelman and J. N. Sturgis, *Biophys. J*., 2003, **85**, 3097–3105.
46. J. D., Lear, H. Gratkowski and W. F. DeGrado, *Biochem. Soc. Trans*., 2001, **29**, 559–564.
47. D. Noy, J. R. Calhoun and J. D. Lear, *Anal. Biochem*., 2003, **320**, 185–192.
48. A. Musatov and N. C. Robinson, *Biochemistry*, 1994, **33**, 13005–13012.
49. A. Musatov, J. Ortega-Lopez, B. Demeler, J. P. Osborne, R. B. Gennis and N. C. Robinson, *FEBS Lett*., 1999, **457**, 153–156.
50. W. F. Stafford III, *Anal. Biochem*., 1992, **203**, 295–391.
51. S. H. White, *Protein Sci.,* 2004, **13**, 1948–1949.

QUERY FORM

Royal Society of Chemistry

Analytical Ultracentrifugation

JOURNAL TITLE:	**RSC-AUC**
ARTICLE NO:	**CH019**

Queries and / or remarks

Query No	Details required	Author's response
AQ1	Please update the reference no.23.	
AQ2	Please update the reference no.24.	